Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

ALT & AST prediction using optimized GAN model based on diabetes and metabolic function data of type 2 diabetes mellitus person

C.Iyyappan[1], R. Latha[2]

[1]*Corresponding Author, Department of computer science, St. Peter's Institute of Higher Education and Research, Chennai, India. Email.id: iyyaravi18@gmail.com*

[2] *Department of computer science and Applications, St. Peter's Institute of Higher Education and Research, Chennai, India. Email id: latharamavel@gmail.com*

ARTICLE INFO

Received: 29 Dec 2024

Revised: 12 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

The incidence of Type 2 Diabetes Mellitus causes metabolic dysfunction and severe health complications. In this paper, optimized machine learning approach predicts liver enzyme levels and enhance diabetes management. The proposed pre-processing algorithms are (i) semi graph theory (SGT) based relation data extraction (ii) Generalized Rough set theory (GRST) for redundant data reduction. The proposed liver enzyme prediction algorithm is (i) Bayesian optimised Sinusoidal regression (BO-SR) (ii) Genetic Algorithm optimised Polynomial Regression (GA-PR). The proposed method applied in the clinical data such as demographics, metabolic function, and sleep data. Semi-graph theory is used for relational structuring and generalized rough set theory is used for eliminating redundant information. Adam optimised Generative Adversarial Networks are used to generate synthetic Thermic Effect of Food data and improve data size and diversity. Genetic Algorithm and Bayesian Optimization ensure optimal parameter selection and improves prediction accuracy. The proposed method has an accuracy of 97.3% in predicting Aspartate Aminotransferase (ALT) and Alanine Aminotransferase (AST) levels, this ensures early diagnosis and intervention for diabetes-related liver complications.

Keywords: Type 2 Diabetes Mellitus, Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Diabetic Management, Liver Enzyme Prediction.

INTRODUCTION

A rising global health concern, primary sign of T2DM includes elevated blood sugar, insulin resistance, and lack of insulin. The increasing prevalence of T2DM worldwide necessitates urgent public health and clinical preventive measures [1]. Urbanisation, ageing populations, and changes in lifestyle choices share a part in the global rise in T2DM. In order to avoid serious health and financial repercussions certain organisations have emphasised how urgent it is to confront this epidemic. An organisation estimates that there were more than five hundred million persons with diabetes in recent past years, and that figure would rise more in next ten years T2DM has more than 90% [2]. The rapid growth is largely driven by increasing obesity rates, sedentary lifestyles, and dietary changes, particularly in developing nations.

Countries like China and India have witnessed a dramatic rise in T2DM cases due to rapid urbanization and dietary shifts towards processed foods. China alone has over 140 million diabetics, the highest in the world. Due to a combination of genetic predisposition and high obesity rates, the MENA has high rates of diabetes. T2DM is highly prevalent in North America and Europe, and better disease management is made possible by increased awareness and access to healthcare. Africa is the fastest-growing region in terms of T2DM cases, where urbanization and changes in lifestyle have led to rising numbers, but healthcare infrastructure struggles to maintain [3]. So, it is important to understand the relation between T2DM and metabolic function. T2DM is an ongoing metabolic disorder characterised by lack of insulin and diminished insulin secretion, resulting in hyperglycemia. The condition is linked to various metabolic functions like glucose homeostasis, lipid metabolism, and energy balance. T2DM primarily

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

develops due to insulin resistance, a condition where peripheral tissues, particularly muscle, liver, and adipose tissue, fail to respond effectively to insulin. It leads to reduced glucose intake in muscle cells, causing decreased energy production, increased hepatic glucose production, aggravating hyperglycemia reducing insulin secretion overtime Metabolic dysfunction in T2DMencompasses not only glucose metabolism but also lipid metabolism. Increased lipo lysis in ae tissue results in increased FFA, exacerbating insulin resistance. Elevated triglycerides and decreased HDL cholesterol are hallmarks of dyslipidaemia, which raises the risk of cardiovascular disease [4]. Additionally, T2DM is linked to compromised mitochondrial activity, which lowers ATP synthesis and raises oxidative stress. It leads to metabolic inflexibility, in which cells find it difficult to effectively transition between the metabolism of fatty acids and glucose. T2DM is characterised by chronic low-grade inflammation, where insulin signalling is disrupted by cytokines such TNF-α and IL-6. Alterations in hormones like leptin and adiponectin disrupt appetite regulation and energy balance [5]. It is necessary to study the relation between T2DM and diabetic function to prevent it.

Insulin secretion, glucose metabolism, and general endocrine regulation are among the metabolic processes that are directly affected by T2DM. The term "diabetic function" describes secretion and usage of insulin by human body. Insulin resistance hampers these processes in T2DM, resulting in persistent hyperglycemia and related consequences. The secretion and action of insulin, a hormone made by pancreatic β-cells that controls glucose uptake, are key functions of diabetes. β-Cell Dysfunction causes pancreatic β-cells failure to compensate for insulin resistance, leading to decreased insulin secretion [6]. Important diabetic function is maintaining glucose homeostasis, which is disrupted in T2DM. The liver continues to produce glucose even when insulin levels are high, worsening hyperglycemia leading to increased hepatic glucose production. Reduced Glucose absorption by muscles and fat cells from the bloodstream also leads to high blood sugar levels.

T2DM affects various hormones that regulate diabetic function which includes glucagon causing increased secretion from pancreatic α-cells leads to excessive glucose production. Amylin is co-secreted with insulin helps regulate glucose levels and appetite, but its function is impaired in T2DM. Hormones like GLP-1 and GIP, stimulates insulin secretion, are less effective in T2DM because of incretin dysfunction. When untreated, T2DM leads to chronic complications that further impair diabetic function. Progressive Loss of β-Cells reduces insulin production. Diabetic Complications increases risk of neuropathy, nephropathy, and retinopathy. Cardiovascular disease is caused due to metabolic and vascular dysfunction [7] Sleep and T2DM are connected proof shows that inadequate sleep length, quality, and patterns have a major impact on diabetes management and risk. Metrics pertaining to sleep duration, efficiency, and disruptions offer important insights into the reciprocal relationship between sleep and type 2 diabetes. Sleep is crucial for metabolic balance, and sleep disturbances can exacerbate insulin resistance and impair glucose management. Less than six hours of sleep every night is linked to higher blood sugar. Long sleep greater than 9 hours per night is linked to metabolic dysregulation, potentially due to underlying health conditions or poor sleep quality. Frequent awakenings or low sleep efficiency disrupts circadian rhythms, impairing glucose metabolism. Poor deep sleep is associated with reduced insulin sensitivity. A prevalent sleep condition in people with type 2 diabetes is obstructive sleep apnoea (OSA), which causes intermittent hypoxia and raises inflammation and oxidative stress. Insulin sensitivity has been seen to improve in patients after CPAP therapy [8].

T2DM itself contributes to poor sleep quality through various physiological and neurological mechanisms. High blood sugar levels at night can lead to frequent urination (nocturia), causing sleep interruptions. Poor control is linked with rising nighttime awakenings and reduced sleep efficiency. Diabetic neuropathy causes pain and discomfort, leading to difficulty falling and staying asleep. Neuropathy and Restless Leg Syndrome (RLS) more common in T2DM, disrupts sleep due to uncontrollable leg movements. The body's internal clock and sleep-wake cycle are impacted by the decreased melatonin secretion associated with type 2 diabetes. Because of their inconsistent sleep habits, shift workers with type 2 diabetes frequently have impaired glycaemic control [9]. Two liver enzymes that are frequently examined to evaluate liver function are AST and ALT. It is essential to keep an eye on the ALT and AST levels of individuals with type 2 diabetes since they are at a higher risk of developing liver-related issues, such as NAFLD and metabolic dysfunction. Since excessive liver fat deposition are common in people with T2DM, elevated ALT and AST readings may indicate underlying liver disease, such as NAFLD. Liver dysfunction is intimately linked to type 2 diabetes. The reduced melatonin release linked to diabetes affects the body's internal clock and sleep-wake cycle. Shift workers with type 2 diabetes often have poor glycaemic control due to irregular sleep patterns [9].

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AST and ALT are two liver enzymes that are regularly tested to assess liver function. It is essential to keep an eye on their ALT and AST levels since those who have these conditions are more likely to experience liver-related issues, such as NAFLD and metabolic dysfunction.

The liver plays a crucial role in maintaining glucose homeostasis. Hyperglycemia is made worse by the liver's overproduction of glucose as a result of insulin resistance. Higher ALT levels are linked to control, which is characterised by elevated fasting glucose and HbA1c values. Liver fibrosis, which is characterised by consistently elevated ALT and AST levels, can lead to cirrhosis. The AST/ALT ratio indicates severe fibrosis or cirrhosis when it is greater than 1, and common NAFLD when it is less than 1. Cardiovascular diseases (CVD) occur in T2DM patients with raised ALT and AST levels. Chronic inflammation, hypertension, and dyslipidaemia are all exacerbated by liver impairment and raise the risk of CVD [11].

Research Gap

Existing research have focused on the Indian population, analysing liver function tests (LFTs) in relation to diabetes [11]. However, these findings are not generalizable to other populations due to genetic, dietary, and environmental differences. Comparative studies across diverse ethnic groups are needed to better understand the influence of demographic factors on liver function parameters in T2DM. Research has highlighted nonlinear relationship between AST, ALT levels and the incidence of T2DM, with focus on single parameter [12]. There is a lack of research on how liver enzyme abnormalities can be used for early diagnosis, risk stratification, and therapeutic interventions in T2DM management and progression.

Problem Statement

T2DM is a progressive illness represented by hyperglycemia and insulin resistance. T2DM impacts an individual's health leading to complications like liver dysfunction. The amount of AST, ALT in diabetic patients should be accurately predicted to avoid severe complications. However, the complexity of T2DM progression, influenced by metabolic functions, demographic factors, clinical indicators, and lifestyle data, poses challenges for accurate prediction and tailored interventions. Traditional methods fail to efficiently process and analyse large datasets with both periodic and non-linear trends. Effective management of T2DM requires personalized approaches to prevent complications and optimize health outcomes.

Contributions

1. To improve data for diabetes management, Adam optimised GAN is used to generate synthetic Thermic Effect of Food (TEF) data. Optimised GAN-generated synthetic data compensates for the limited availability of real medical data, particularly for Thermic Effect of Food (TEF) in diabetic individuals.
2. To improve structural connectivity for relational data representation Semi graph theory is used. Semi-Graph Theory preserves the interconnected data, making it useful for correlation-based analyses in diabetes management.
3. To eliminate redundant data and improve the performance generalized rough set theory is used. Removing irrelevant data reduces complexity, improving prediction accuracy for ALT and AST enzyme levels.
4. To capture periodic variations in metabolic and sleep data that influence liver enzyme fluctuations Sinusoidal regression is used and to identify non-linear relationships in metabolic function and TEF, liver enzymes, Polynomial regression is used, improving prediction accuracy and generalization across diverse patient profiles.
5. To improve the prediction accuracy of liver enzymes Genetic Algorithm and Bayesian optimization algorithms are used. Fine-tuning the parameters of machine learning algorithms for more reliable liver enzyme predictions.

LITERATURE SURVEY

An important biomarker for comprehending a number of metabolic diseases, especially T2DM its consequences, is the AST/ALT ratio. The results of contemporary research examining the connection between the AST/ALT ratio and

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

T2DM, gestational diabetes mellitus, diabetic retinopathy, cardiovascular mortality, and other associated disorders are presented in this review of the literature. Numerous metabolic and cardiovascular disorders, such as T2DM, GDM, hypertension, sepsis, coronary artery disease, have been extensively researched using the AST-to-ALT ratio as a biomarker. The review that follows summarises the body of research on the connection between these health outcomes and the AST/ALT ratio. An investigation on the nonlinear link between the occurrence of T2DM and the AST/ALT ratio was carried out by the author in [12]. found that variations in AST, ALT detects disease. The findings highlight the significance of liver enzyme ratios in metabolic disorders, suggesting that abnormal AST/ALT levels might reflect underlying insulin resistance and hepatic dysfunction TG/HDL-C and the AST/ALT ratio have been compared by the author as risk factors for GDM in [13].

It was concluded that AST/ALT serves as a more reliable independent risk factor emphasizing the role of hepatic function in gestational metabolic disturbances and suggests that liver enzyme monitoring during pregnancy could aid in early GDM risk assessment. The correlation of AST/ALT, cardiovascular and results in death in hypertensive individuals has been investigated by the author in [14]. The results support the idea that liver function abnormalities contribute to cardiovascular problems in hypertensive individuals by showing a correlation between raised AST/ALT ratio and increased mortality risks. In [15] the potential value of AST/ALT as a predictive biomarker in critically ill patients is demonstrated by the finding that a higher ratio was linked to worse clinical outcomes. It was recognised that liver failure had a significant role in the development of sepsis and the outcome for patients. The effect of AST/ALT levels on all-cause mortality in patients with stable heart disease was discussed by the author in [16]. Higher death rates were associated with an elevated ratio, according to the secondary analysis, further reinforcing the potential of this biomarker in cardiovascular risk stratification.

Author has evaluated prognostic significance of ratio in comparison to bilirubin [17]. The data revealed ratio was a superior detector of outcomes compared to bilirubin levels, underlining its therapeutic value in managing critically unwell cardiac patients. Potential of the elevated AST/ALT ratio as a marker for diabetes complications was highlighted by the author's discovery in [18] that it was linked to type 2 diabetic peripheral neuropathy in a Chinese population. The relationship between AST/ALT and death in critically unwell congestive cardiac patients has been investigated by the author, showing that elevated levels correlated with worse clinical outcomes [19]. In [20] Author has noted that the AST/ALT ratio was a useful indicator of functional severity in chronic heart failure with reduced left ventricular ejection fraction. Author has examined the non-linear association between AST/ALT and mortality in critically ill older patients, suggesting that deviations from an optimal range could indicate higher mortality risks [21]. The findings contribute to the growing body of evidence supporting AST/ALT as a useful marker in elderly patient populations. However, the study had certain limitations, including its retrospective design, which may introduce bias, and the lack of consideration for confounding factors such as comorbidities and medication use.

In a cross-sectional investigation, the authors looked at the AST/ALT ratio's function as a stand-alone risk factor for diabetes retinopathy in [22]. According to their findings, there may be a connection between liver malfunction and microvascular problems because a greater AST/ALT ratio was linked to an increased risk of diabetic retinopathy. The cross-sectional form of the study, however, makes it impossible to demonstrate causation, and more longitudinal research is needed to validate the results. In [23] study found that an elevated ratio correlated with liver fibrosis and disease progression, making it a useful clinical marker. However, this study, being over two decades old, may not fully reflect current diagnostic advancements. Furthermore, it focused solely on hepatitis C, limiting its applicability to other liver conditions. The author has investigated the connection between the severity of chronic heart failure 6[24] and the AST/ALT ratio. According to their research, patients with a lower left ventricular ejection fraction were expected to have worse functional status if their AST/ALT ratio was higher. Notwithstanding these encouraging findings, the study lacked a thorough evaluation of potential confounding factors, including systemic inflammation and kidney function, and had a small sample size. The study in [25] reported that an increased transferase ratio was linked to higher risk of adverse events making a potential prognostic biomarker. Author in [26] examined pregnancy cholestasis, a disorder that usually manifests in the third trimester of pregnancy. The study examined liver enzyme levels and their relationship to pregnancy outcomes using retrospective cohort methods. According to their research, ratio can be sign of a higher chance of unfavourable consequences for both the mother and the foetus. The study did

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

have some drawbacks, though, such as a small sample size and no control for other liver problems, which would have limited how broadly the findings can be applied. Table 1 provides the existing methods in ALT& AST measurement

Table 1 Existing methods in ALT& AST measurement.

Reference	Study	Key Findings	Methods Used	Drawbacks
[12]	Nonlinear Association Between T2DM and the AST/ALT Ratio	- Nonlinear relationship identified; higher ratios associated with lower T2DM incidence in prediabetic individuals.	- Follow-up study, cohort analysis	- Limited generalizability due to specific population
[13]	Alt and AST are linked to the risk of gestational diabetes	- An increased AST/ALT ratio is linked to a higher risk of gestational diabetes.	- Case-control study	- Potential confounding factors not fully controlled
[14]	Mortality in H and the Ratio of ALT, AST	- In hypertensive patients, a increased ratio of AST, ALT is linked to higher cardiovascular and all-cause mortality.	- Cohort study, retrospective analysis	- Single-center study may limit external validity
[15]	Ratio of AST/ALT in Sepsis	AST/ALT ratio has been shown to have diagnostic and prognostic value in patients with sepsis and septic shock.	- Observational study	- limited sample size; findings might not be generally relevant
[16]	The AST/ALT Ratio and Cardiovascular Disease	- An increased AST/ALT ratio is linked to all-cause mortality in patients with stable coronary artery disease.	- Retrospective cohort study	- Data derived from medical records may lack completeness
[17]	Bilirubin vs the AST, ALT levels in Cardiogenic Shock	- AST, ALT ratio prognostic value in relation to bilirubin in cardiogenic shock patients.	- Comparative analysis	- Variability in clinical management across case
[18]	DPN	-An increased risk of diabetic peripheral neuropathy is linked to higher AST/ALT ratios (OR: 2.413).	- Cross-sectional study	- Single-center study may limit generalizability
[19]	Death	-. An increased AST, ALT levels is linked to a higher risk of death in critically unwell congestive heart failure patients.	- Cohort analysis	- Limited demographic diversity
[20]	Chronic Heart Failure	- In individuals with chronic heart failure, higher AST, ALT levels were associated with worse functional severity.	- Prospective cross-sectional study	- Small sample size may affect statistical power

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Reference	Study	Key Findings	Methods Used	Drawbacks
[21]	Death & Severely Ill Elderly persons	The AST, ALT levels, death have connection; thresholds where risk dramatically increases have been found.	- Retrospective cohort study	- Single-center analysis may limit applicability
[22]	Diabetic Retinopathy Risk Factor Study	The AST, ALT levels and the risk of diabetic retinopathy are closely linked	Cross-sectional study	Single-center study may limit generalizability.
[23]	Hepatitis C Prognosis	AST/ALT ratio provides information about disease severity and prognosis in HCV-related chronic liver disease.	Observational study	Limited in isolation; other factors needed for comprehensive assessment.
[24]	Chronic Heart Failure Severity Prediction & Reduced LVEF & Functional severity prediction possible.	A high AST/ALT ratio predicts the functional severity of chronic heart failure with reduced LVEF.	Observational study	May not establish causation.
[25]	Unstable Angina & Cardio-Cerebral Events Association & Increased transferase ratios are associated	In unstable angina, an elevated transferase ratio (AST/ALT) causes unfavourable cardio-cerebral events.	Study of retrospective group	Potential for selection bias and residual confounding.
[26]	Clinical Significance in Third Trimester Pregnancy	Highlights the clinical significance of ICP typically occurring in the third trimester, emphasizing its impact on pregnancy outcomes due to increased risk of preterm labor, fetal distress, etc., necessitating close monitoring and appropriate management strategies.	retrospective cohort study methods	limited sample size, which restricts how far the findings can be applied.

Table 1 summarizes critical aspects discussed in the literature review, providing an overview of key findings, methods used, and limitations of each study about the influence of the AST, ALT on T2DM and other health complications. According to the research, the ratio detects metabolic health, specifically of type 2 Diabetes and its effects. Its significance in clinical practice for early diagnosis and intervention techniques is highlighted by its nonlinear relationship with the incidence of diabetes and relationship with other health issues. The results of several clinical research are combined in this review to provide a understanding of AST/ALT ratio in connection to diabetes and its complications indicating its use in clinical assessment and intervention techniques.

METHODOLOGY

Figure 1 displays the block diagram for the proposed methodology.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

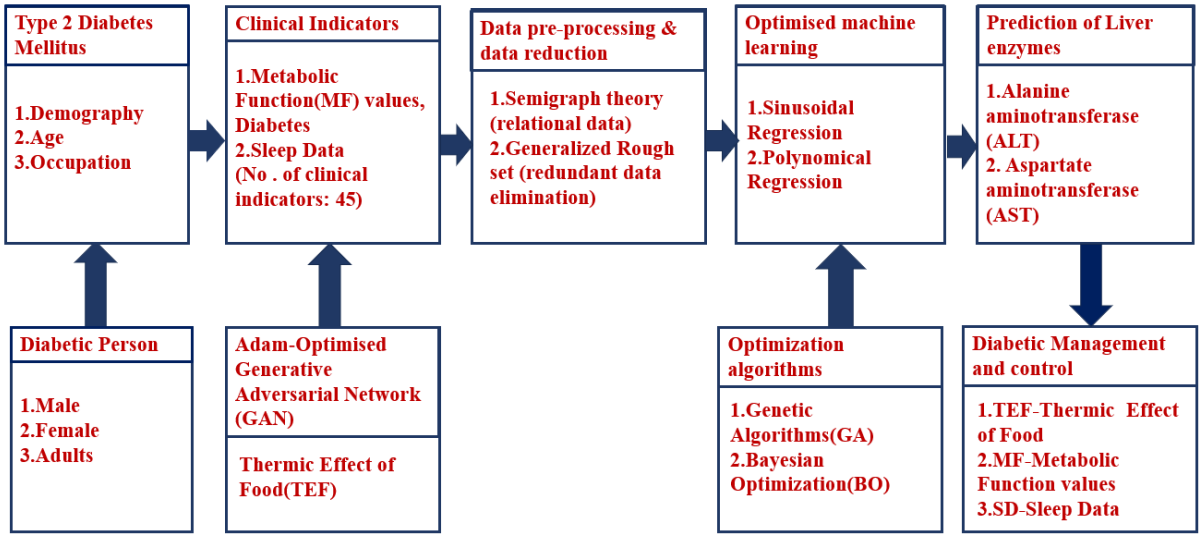


Figure 1 Block diagram for predicting liver enzymes AST, ALT

Diabetic Person

Individuals affected by Type 2 Diabetes Mellitus are classified into three groups male, female and adults. The classification is essential as gender and age differences influence the effect of diabetes on each group. Men have more chances of insulin resistance because of high visceral fat accumulation. Women suffer from hormonal fluctuations affecting blood sugar control and gestational diabetes increases the chances of developing T2DM later in life. Adult persons over the age of 40 are the most affected by T2DM due to age-related insulin resistance, and younger adults are at an increasing risk due to sedentary lifestyle and obesity. This classification is required for the personalized diabetes management strategies based on age and gender to improve early detection, treatment and long-term diabetes care.

Type 2 Diabetes Mellitus

Demography

Demographic factors determine an individual's risk of developing type2 Diabetes Mellitus. Demographic features like genetics, race, geographic location influence the development and progression of diabetes in an individual. Geographic location where the people live influence the risk of diabetes due to food habits, lifestyle and access to healthcare. Certain ethnic groups have a higher predisposition to developing Type 2 Diabetes.

Age

Age is an important threat element for T2DM. Adults are more likely to be affected, especially those over 45. T2DM is common among teenagers, young adults, because of obesity and sedentary lifestyles. Age-related insulin resistance and reduced pancreatic function contribute to diabetes development in older adults.

Occupation

A person's job occupation impacts the risk of diabetes. Sedentary jobs increase the risk due to the lack of physical activity and dietary habits. Jobs involving physical labour has lower risk but shift work causes irregular eating and sleeping patterns resulting in metabolic disorders. Figure 2 shows the distribution of diabetic persons.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

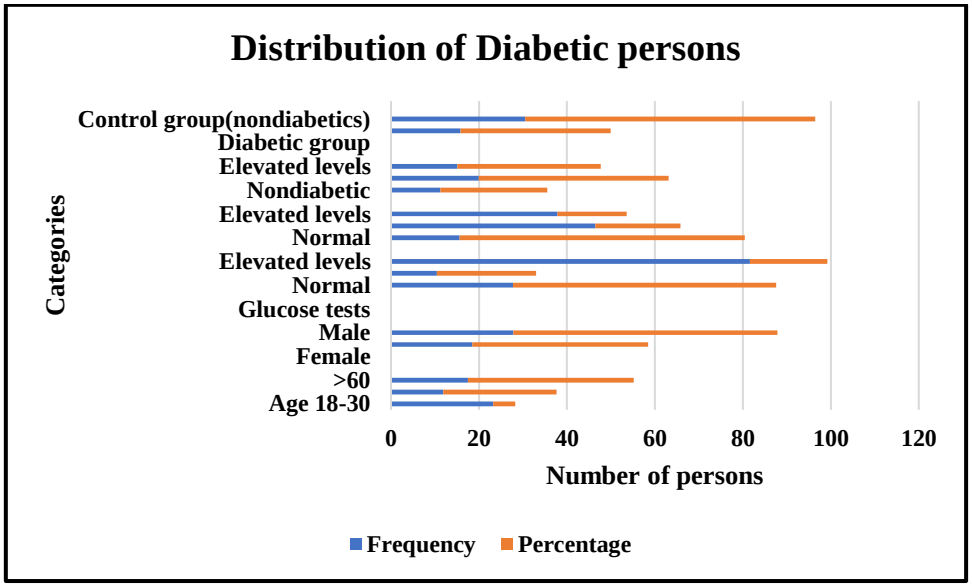


Figure 2 Distribution of Diabetic persons in various categories

Distribution of Diabetic persons in various categories like Control group (nondiabetics), Diabetic group. Elevated levels, Normal in Glucose, gender based and age based is shown in Figure 2.

Adam optimised GAN -Adaptive Moment Estimation optimised Generative Adversarial Network

Adam optimised GAN is used to generate synthetic TEF data as Thermic Effect of Food (TEF) data is limited in real world. The Adam optimizer enhances the training process of GANs by providing faster convergence. It improves the stability of GAN for generating high-quality synthetic TEF data without limited patterns. The dataset referenced in [27] includes 502 food items with detailed nutritional attributes like glycemic index, calorie content, macronutrients (carbohydrates, protein, fat), diabetes and blood pressure suitability, and mineral contents like sodium, potassium, magnesium, calcium, and fiber. TEF is calculated for each food by the given formulas and the new TEF dataset is used as input for Adam optimised GAN. Table 2 explains the function of Adam optimised GAN in generating the synthetic TEF dataset.

$$TEF = Total\ Energy\ Intake * TEF\ Percentage \quad (1)$$

$$TEF = (Calories\ from\ Carbohydrates * 0.75) + (Calories\ from\ Protein * 0.25) + Calories\ from\ fat * 0.025) \quad (2)$$

Table 2 Process Flow for Adam-Optimized GAN in TEF Data Generation

Process	Function of Adam-optimised GAN
Data collection	Gathers real-world food intake data based on TEF for training.
Generator	Creates synthetic TEF values by learning complex data distributions using adaptive moment estimation (Adam) for efficient weight updates.
Discriminator	Differentiates real TEF dataset from generated synthetic TEF data, leveraging Adam optimizer for better convergence stability.
Synthetic Data Generation	GAN generates artificial TEF data to expand the dataset, reducing data scarcity and maintaining realistic patterns.
Data Augmentation	Improves dataset variability and ensures better generalization in machine learning models by optimizing GAN's weight updates through Adam

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Feature Learning	Enhances the prediction of hidden patterns in TEF-based food data, improving the quality of synthetic data and reducing bias.
Improving Model Robustness	Provides additional training samples with an optimized GAN, reducing overfitting through better weight updates and convergence control.
Diabetes Management Insights	Enhances TEF prediction accuracy for better diabetes control strategies by using a stable and well-optimized GAN model.
Benefits of Using GANs	1.Increases data availability by generating new TEF data to improve generalization 2. Overfitting is reduced by allowing Sinusoidal and polynomial regression to not depend on limited real-world dataset. 3. Optimized GAN based TEF dataset leads to better accuracy in liver enzyme predictions, contributing to improved diabetes management.

Table 2 shows that generated dataset is augmented by Adam-Optimised GAN for better training of machine learning models. It ensures that Optimised GAN learns dietary patterns to generate realistic TEF data. Table 3 provides the pseudocode for Adam-optimised GAN-Based TEF Data Processing and Figure 3 shows the Adam-optimised GAN generated TEF data.

Table 3 Pseudocode for Adam-optimised GAN-Based TEF Data Generation

```
1. **Initialize Environment**
  - Clear workspace, close all figures, and reset command window.
2. **Load and Preprocess Dataset**
  - Load the TEF dataset from a CSV file.
  - Display column names for verification.
  - Extract numerical features and handle missing values.
  - Normalize numerical features for GAN training.
3. **Define GAN Parameters**
  - Set dimensions for numerical features.
  - Define latent space size for random noise input.
  - Set hyperparameters such as learning rate, batch size, and number of epochs.
4. **Define Generator Network**
  - Input: Random noise vector.
  - Layers:
    - Fully connected layers with ReLU activation.
    - Output layer with Tanh activation to match normalized feature range.
  - Output: Synthetic TEF data.
5. **Define Discriminator Network**
  - Input: Real/synthetic TEF data.
  - Layers:
    - Fully connected layers with Leaky ReLU activation.
    - Output layer with Sigmoid activation to classify real vs. fake data.
  - Output: Probability score indicating whether the input is real or fake.
6. **Define Adam Optimizer**
  - Set learning rate (e.g., 0.001).
  - Set beta1 and beta2 for first and second moment estimates (e.g., 0.9 and 0.999).
  - Set epsilon for numerical stability (e.g., 1e-7).
7. **Train the GAN with Adam Optimizer**
  - For each epoch:
    1. Sample a mini-batch of real TEF data.
    2. Generate random noise.
    3. Generate synthetic TEF data using the generator.
```

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4. Compute gradients for discriminator and generator using Adam optimizer.
5. Update discriminator weights:
   - Compute loss on real data (real label = 1).
   - Compute loss on synthetic data (fake label = 0).
   - Update weights using Adam update rule.
6. Update generator weights:
   - Compute loss by feeding synthetic data through discriminator (real label = 1).
   - Update weights using Adam update rule.
7. Display training progress every few epochs (e.g., loss values).

8. **Generate Synthetic TEF Data**
   - Define the number of synthetic samples to generate.
   - Generate random noise vectors.
   - Use the trained generator to produce synthetic TEF data.
   - Denormalize synthetic data to match original scale.
9. **Save Synthetic Data**
   - Convert synthetic data into a table format with appropriate column names.
   - Save the table as a CSV file.
10. **Performance Metrics (Optional)**
   - Evaluate the quality of generated data using metrics such as:
     - Mean Absolute Error (MAE) between real and synthetic data samples (if applicable).
     - Mean Squared Error (MSE) between real and synthetic data samples (if applicable).
     - Root Mean Squared Error (RMSE) between real and synthetic data samples (if applicable).
     - Coefficient of Determination (R^2) to assess model fit.
   - Display performance metrics.
11. **Display Results**
   - Print final performance metrics (e.g., MAE, MSE, RMSE, R^2).

synthetic_food_data.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
								syntheticfooddata								
	GlycemicIn...	Calories	Carbohydra...	Protein	Fat	SuitableFor...	SuitableFor...	SodiumCon...	PotassiumC...	Magnesium...	CalciumCon...	FiberContent	CaloriesFro...	CaloriesFro...	CaloriesFro...	TEF_kcal_
	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
1	GlycemicIn...	Calories	Carbohydr...	Protein	Fat	SuitableFor...	SuitableFor...	SodiumCo...	Potassium...	Magnesiu...	CalciumCo...	FiberContent	CaloriesFro...	CaloriesFro...	CaloriesFro...	TEF_kcal_
2	1.806113	4.764338	1.254458	0.7876084	3.720567	4.729177	-0.120883	-0.5691848	0.1648991	2.476404	1.398481	0.3209796	4.643791	1.081572	4.36384	3.724791
3	1.660192	4.349265	2.832918	-0.3057942	1.873217	5.586344	3.410877	0.2323978	3.382688	3.201291	0.9249661	1.857153	1.263127	-1.764453	2.287385	2.739055
4	0.7665294	5.955522	5.639191	-3.293538	-0.0792346	4.103182	-0.3353262	1.46182	3.752684	-0.664444	-0.3804486	-2.670279	3.5294	-1.430237	2.688838	5.764447
5	2.891364	6.146658	4.138903	-3.812251	1.269291	3.410605	1.033885	3.034034	-1.382485	0.8871453	-2.163261	1.141635	4.064872	-0.02056503	6.082044	4.276321
6	1.942482	4.120525	1.638996	1.634017	3.470693	1.81395	1.884181	-0.1440206	1.178697	1.073404	1.875346	0.8614231	3.598937	-0.2255311	3.793506	3.542326
7	2.081495	7.088074	2.071004	0.4649482	2.172268	6.920381	0.3405758	-1.933608	5.342792	-0.3680453	1.379648	1.099466	4.82181	-1.352994	1.757308	1.392308
8	0.7783085	4.216335	4.175868	0.1147232	2.445853	3.200921	1.630612	1.976361	3.096864	2.935393	-0.1626263	0.003576517	4.265765	-0.5440307	1.616824	3.288908
9	-0.6990919	5.249491	3.028739	0.1762257	5.56101	5.968773	2.030905	3.009729	-1.456282	4.836058	-0.4389668	2.707673	2.640643	-0.1328976	4.252731	4.589387
0	1.1987	4.229227	0.8866071	1.700576	3.017441	4.246982	1.996775	1.521516	1.769708	1.149068	1.453852	1.632316	3.618444	0.641492	3.538736	2.219401
1	1.526209	7.194117	5.122813	0.4609468	-0.1731458	5.788181	2.058275	1.377254	5.812348	1.775525	0.04465628	2.273275	0.5683299	-1.537885	2.359577	2.789212

Figure 3 Synthetic TEF data generated by Adam-optimised GAN.

Dataset

Dataset includes metabolic function, diabetes, sleep data. The datasets are essential to understand the influence of diabetes in ALT, AST levels, the indicators of liver health. Each dataset contributes features that help in accurate prediction of ALT, AST.

Metabolic Function and diabetes dataset

The body's capacity to process energy, control blood sugar, and ensure functioning of organs is referred to as metabolic function. It is related with liver health and overall metabolic health. It is useful to evaluate the effect diabetes on liver as Insulin resistance and other metabolic diseases cause NAFLD, prevalent among diabetics. Data like glucose levels, insulin resistance, and lipid profiles, provide information about the progression of diabetes and its complications.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Sleep Data

Sleep is important in metabolism, insulin sensitivity, and overall diabetes management. Poor sleep quality can worse glucose control and increase liver enzyme levels. Insufficient sleep leads to metabolic disturbances, increasing insulin resistance and worsening liver enzyme levels. OSA leads to inflammation and liver stress. Circadian rhythm misalignment affects glucose metabolism and may alter ALT/AST patterns. Table 4 provides features present in the datasets.

Table 4 Metabolic and Sleep Features Influencing AST & ALT Levels in Diabetes patients

Feature	Description	Importance
Blood Glucose Levels	Fasting, and HbA1c values	High glucose levels indicate poor metabolic control, impacting liver enzymes.
Lipid Profile	Cholesterol, HDL, LDL, Triglycerides	Dyslipidemia is common in diabetes and affects liver function.
Height, Weight	Height and weight ratio	High BMI correlates with fatty liver disease and abnormal ALT/AST.
Blood Pressure	Systolic and Diastolic	Hypertension is associated with metabolic syndrome and liver dysfunction.
Liver Function Tests	ALT, AST	Used as target variables in prediction.
Diabetes Type	Type 2	Type 2 diabetes has a stronger correlation with metabolic dysfunction and liver issues.
HbA1c Levels	Average blood glucose over 3 months	A higher value indicates poor glucose control, affecting liver function.
Diabetes-Related Complications	Neuropathy, Retinopathy, Nephropathy, etc.	Patients with complications may have higher ALT/AST due to organ stress.
Sleep Duration	Number of hours slept each night in total	Increased ALT/AST is linked to shorter sleep duration.
Sleep Quality	Self-reported sleep efficiency	Poor sleep quality affects glucose metabolism and liver function.

Table 3 shows Metabolic, Diabetes, and Sleep Data together providing an extensive dataset for predicting ALT and AST levels in diabetes patients. It improves accuracy, provides better clinical information, and helps in the early detection of liver dysfunction.

Pre-processing and reduction

Pre-processing and data prepare the data for optimised machine learning algorithms. Semigraph Theory and Generalised Rough Set Theory handles missing values, normalizes data, selects related features, and reduces dimensionality in preprocessing.

Semigraph theory

Semigraph Theory is used to present the relationship and dependencies within the dataset. It identifies the relevant features in metabolic, diabetes dataset, sleep data by analysing their relationships by feature selection. Interdependent features like blood glucose, BMI, ALT, AST, HbA1c, sleep duration is represented as a node in a semi graph and edges between nodes represent their statistical correlation and influence. Features with weak connections to ALT, AST are removed to reduce dimensionality. Influence of different factors like sleep duration, metabolic markers on each other and on AST, ALT levels are understood by dependency analysis. Traditional feature selection methods do not capture interactions between multiple variables. Degree centrality and betweenness centrality identify the important features and the features with highest connectivity to ALT, AST are retained. Table 5 presents the comparison of pre-processing techniques with semi-graph theory.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 5 Performance comparison of other pre-processing techniques with Semigraph theory.

Criteria	Semi-Graph Theory (Proposed method)	Feature Selection (PCA, LASSO)	Clustering-Based Preprocessing	Fuzzy Logic-Based Preprocessing	Statistical Methods (Normalization, Standardization)
Handling Complex Relationships	10	6	5	7	3
Preservation of Structural Information	10	4	3	6	2
Handling Missing Data	9	5	6	7	4
Adaptability to Non-Linear Data	10	4	7	9	3
Scalability for Large Datasets	8	7	5	6	10
Improvement in Prediction Accuracy	10	8	6	8	5
Applicability to Medical Data (Diabetes Complications)	10	7	5	8	4
Robustness Against Outliers	*9*	*5*	*6*	*7*	*3*

Table 5 shows that Semigraph Theory performs better than other techniques as it captures complex dependencies and interactions in a graph structure, handles complex relationships, missing data, and non-linear dependencies in the medical dataset used for predicting AST and ALT levels in diabetic patients. Figure 4 shows the interconnection of features in AST, ALT prediction.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

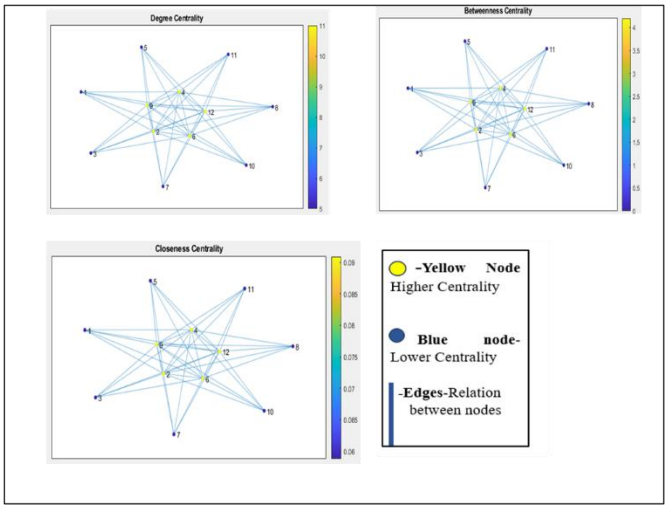
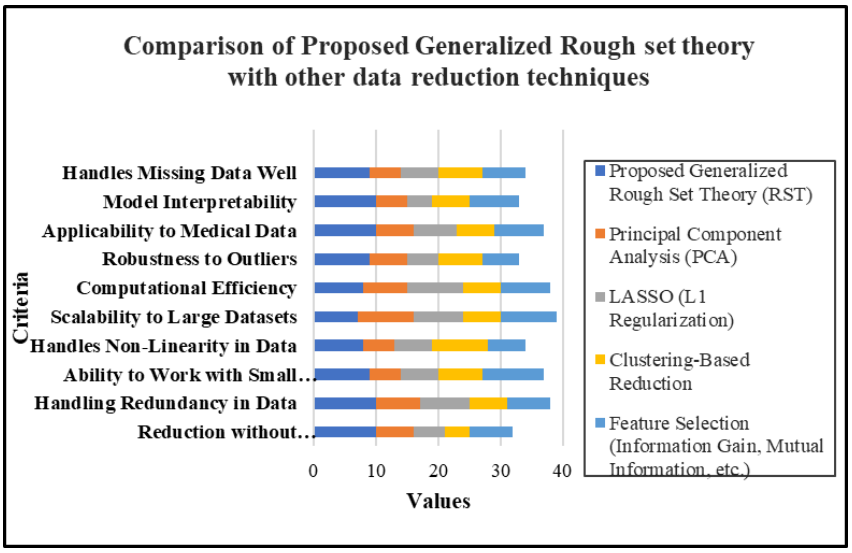


Figure 4 Degree centrality, betweenness centrality and closeness centrality of the features using Semi-graph theory for AST, ALT prediction.

Network's structure and nodes within that network is shown in figure 4 by different centrality metrics. The use of color-coding allows for a quick assessment of which nodes are most central according to each measure. Node 12 has the highest degree centrality; Node 4 has the highest betweenness centrality

Generalized Rough set theory (GRST)

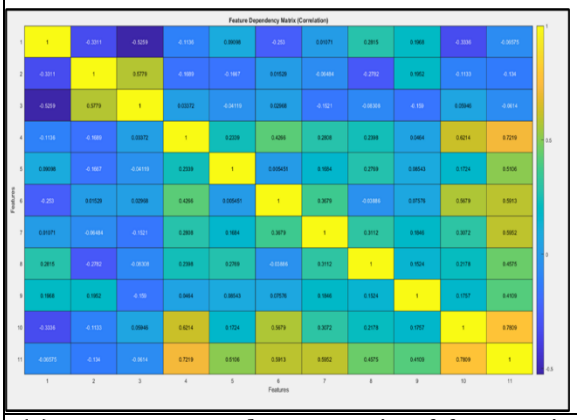
Generalized Rough Set Theory is used to identify and remove irrelevant data without losing essential data to predict liver enzymes AST, ALT. GRST is used because of its ability to handle continuous, uncertain, and missing data in medical datasets efficiently than traditional rough set theory. GRST functions by organising the data like clinical indicators, demographic data, and GAN-generated TEF as attributes. It uses flexible relations to handle mixed data like lower approximation has data definitely belonging to a class, upper approximation possibly belonging to a class, boundary region contains uncertain cases. It measures the importance of each feature in decision-making by attribute significance and removes redundant attributes, keeping only a minimal subset (reduct) for prediction of liver enzymes. Figure 5 shows the performance comparison of generalized rough set theory with other data reduction techniques. The reduced and optimized dataset is then fed into the regression models (Sinusoidal & Polynomial Regression) for AST & ALT prediction.



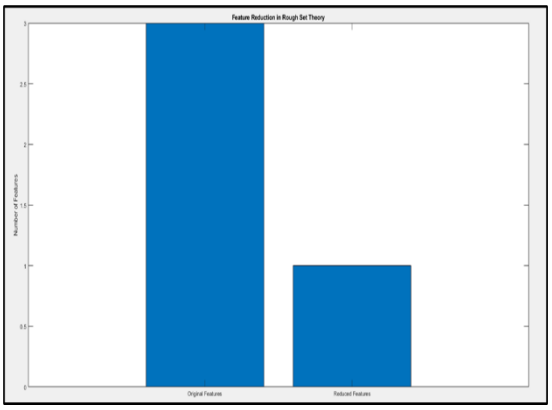
Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 5 Performance comparison of Generalized Rough set theory and other data reduction techniques. Generalized Rough Set Theory performing well in handling missing data, model interpretability, applicability to medical data, and handling redundancy making it suitable for medical applications is shown in figure 5. Figure 6 provides the feature dependency and reduction in predicting liver enzymes to manage diabetes.



(a) Feature Dependency Matrix of features in clinical indicators dataset



(b) Feature reduction using Generalized Rough Set theory

Figure 6 (a)Feature Dependency Matrix of features in clinical indicators dataset(b) Feature reduction using Generalized Rough Set theory

Figure 6(a) presents the connections between different features in a dataset by correlation matrix. A yellow and green cell indicates that the two features tend to increase or decrease together. A blue cell indicates that as one feature increases, the other tends to decrease. Feature reduction by Roughset theory is shown in figure 6b. Bar graph indicates the number of features before and after a feature reduction process using Generalized Rough Set Theory.

Optimized machine learning and optimization algorithm

Pre-processed and reduced dataset with semigraph theory and generalized rough set theory is used for training optimized machine learning algorithms-sinusoidal regression and polynomial regression. The machine learning models are used to develop an accurate predictive model for ALT and AST levels, the key indicators of liver function, affected by Type 2 Diabetes Mellitus. Machine learning is applied to identify the patterns in metabolic function, diabetes, and sleep data that influence ALT and AST levels.

Sinusoidal Regression

Sinusoidal regression analysis is used to analyse the periodic pattern in the dataset. Metabolic functions, sleep cycles, thermic effect of food exhibit periodic variations. Sinusoidal Regression captures periodic variations in ALT/AST influenced by metabolic function, sleep patterns. It has better accuracy for biological data compared to linear models and it is useful for time-series prediction in diabetes-related data. Sinusoidal regression fits a sine wave to the dataset by

$$y = A\sin(Bx + C) + D \qquad (3)$$

Where,

y= Predicted ALT/AST level

x= Time-based or another independent variable

A= Amplitude (occurring variation)

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

B= Frequency (repeating pattern)

C= Phase shift (adjusts the wave along the x-axis)

D= Vertical shift (baseline ALT/AST levels)

This equation captures oscillatory behaviour in ALT/AST fluctuations.

Polynomial Regression

Polynomial regression is used to handle nonlinear trends in the dataset. Liver enzymes ALT and AST exhibit complex, non-linear relationship with metabolic function values and sleep data. Polynomial regression can capture these complex relationships and produce more accurate results than linear regression when relationships are nonlinear. The model fits a polynomial equation of degree n:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n \quad (4)$$

Where,

y = Predicted ALT/AST level

x = Input feature like blood glucose, sleep duration

$a_0, a_1, \ldots, a_n$ =Coefficients learned from data

Optimization Algorithm

The selected machine learning models to minimize error and improve model's performance. The optimization algorithms used here are,

1.Genetic Algorithm (GA)

2.Bayesian Optimization

Genetic Algorithm (GA)

Genetic Algorithm is a heuristic optimization-based natural selection. GA is suitable for optimizing nonlinear and non-convex problems, which are common in medical data due to the intricate relationships between clinical indicators. It tunes the parameters of regression models amplitude, frequency, phase shift to improve prediction accuracy for liver enzymes.

Bayesian Optimization Algorithm

Bayesian optimization uses probabilistic model probabilistic models to explore the parameters of regression models, reducing computational cost compared to other methods. Bayesian Optimization refines model parameters by balancing exploration and exploitation, making better predictions of ALT and AST levels. Table 6 provides the function of regression model optimization techniques and table 7 provides the pseudocode for optimised Polynomial regression and Sinusoidal Regression.

Table 6 Regression model optimization techniques and process involved

Regression Type	Polynomial	Sinusoidal
Optimization Technique	Genetic Algorithm	Bayesian Optimization
Reason for Optimization	To automatically find the optimal polynomial degree and coefficients for best fit, avoiding overfitting or underfitting.	To efficiently determine the best sinusoidal parameters (amplitude, frequency, phase shift) for modelling periodic variations in liver enzyme levels.
Process involved in optimization	1.Initialize Population: Randomly generate polynomial equations with different degrees and coefficients.	1. Define the Objective Function: Minimize the prediction error (MSE) by tuning sinusoidal

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

	2.Fitness Evaluation: Compute the prediction error (MSE/RMSE) for each equation. 3.Selection: Retain the best-performing polynomial models. 4.Crossover & Mutation: Combine the best models and introduce slight variations. 5.Convergence: Repeat until an optimal polynomial equation is found.	parameters. 2. Select a Model for Surrogacy: It tries to estimate the unknown desired function using a probabilistic model like the Gaussian Process (GP). 3. Select an Acquisition Function: Use Expected Improvement (EI), Upper Confidence Bound (UCB), or Probability of Improvement (PI) to decide where to sample next. 4. Evaluate the Objective Function: Test the selected sinusoidal parameters and compute the prediction error. 5. Update the Surrogate Model: Improve the GP model with new data. 6. Iterate Until Convergence: Repeat the steps until an optimal sinusoidal function is found.
Advantages	1.Automatically selects best polynomial degree 2. Prevents overfitting 3. Handles non-linearity in medical data 4. Suitable for short-term metabolic trends	1.Captures periodic/cyclic trends in metabolic function 2. Automatically tunes sinusoidal parameters 3. More efficient than grid search 4. Useful for long-term biological cycles

Table 7 Pseudocode for optimised Polynomial regression and Sinusoidal Regression for AST, ALT prediction.

Genetic Algorithm optimised Polynomial Regression	Bayesian optimised Sinusoidal Regression
1. **Initialize Environment** - Clear workspace, close all figures, and reset command window. 2. **Load and Preprocess Data** - Read Excel file into a table, preserving column names. - Display column names for verification. - Make column names valid for MATLAB. - Dynamically assign feature and target variables: - Feature1: First column. - ALT: Third last column. - AST: Second last column. - ALT/AST Ratio: Last column. - Remove rows with NaN values from all variables. - Normalize Feature1 for numerical stability. 3. **Polynomial Regression with Genetic Algorithm** - Set polynomial degree (e.g., 3) to avoid overfitting. - Define GA options: - Population size (e.g., 100).	1. **Initialize Environment** - Clear workspace, close all figures, and reset command window. 2. **Load Dataset** - Define file path for the dataset. - Detect import options to preserve column names. - Read the dataset into a table. - Display column names for verification. 3. **Extract Input and Output Variables** - Extract input features (X) from the first 7 columns. - Identify and extract the output variable (Y), specifically 'Predicted AST/ALT'. - Handle missing column by throwing an error if not found. 4. **Define Objective Function for Bayesian Optimization**

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Maximum generations (e.g., 200).
- Set bounds for coefficients to prevent large oscillations:
  - Lower bound (e.g., -1000).
  - Upper bound (e.g., 1000).
- Record start time for computational time measurement.
- Perform GA optimization for each target variable:
  - ALT:
    - Define objective function: Sum of squared differences between predicted and actual ALT.
    - Run GA to find best coefficients.
    - Predict ALT using best coefficients.
  - AST:
    - Define objective function: Sum of squared differences between predicted and actual AST.
    - Run GA to find best coefficients.
    - Predict AST using best coefficients.
  - ALT/AST Ratio:
    - Define objective function: Sum of squared differences between predicted and actual ALT/AST Ratio.
    - Run GA to find best coefficients.
    - Predict ALT/AST Ratio using best coefficients.
- Record end time and calculate computational time.

4. **Performance Metrics**
- Calculate metrics for each prediction:
  - **RMSE (Root Mean Squared Error)**: sqrt(mean((Actual - Predicted)^2)).
  - **MAE (Mean Absolute Error)**: mean(abs(Actual - Predicted)).
  - **MSE (Mean Squared Error)**: mean((Actual - Predicted)^2).
  - **R^2 (Coefficient of Determination)**: 1 - sum((Actual - Predicted)^2) / sum((Actual - mean(Actual))^2).
- Calculate computational time in seconds.

5. **Visualization**
- Create a figure with three subplots:
  - Plot actual vs predicted values for ALT, AST, and ALT/AST Ratio.
  - Use a smooth curve for predicted values.
  - Include titles, legends, labels, and grid for clarity.

6. **Display Results**
- Print performance metrics for each prediction:
  - RMSE.
  - MAE.
  - MSE.
  - R^2.
- Print computational time in seconds.

- Create a function that computes the mean squared error (MSE) between predicted and actual values using a sinusoidal regression model.
- The model includes parameters for sinusoidal components and linear terms.

5. **Define Optimization Variables**
- Define bounds for each parameter in the sinusoidal regression model:
  - A1, A2: Amplitudes.
  - B1, B2: Frequencies.
  - C1, C2: Phases.
  - P1, P2: Linear coefficients.
  - D: Constant offset.

6. **Run Bayesian Optimization**
- Set up Bayesian optimization with the defined objective function and variables.
- Specify optimization settings:
  - Maximum number of objective evaluations.
  - Acquisition function for optimization.
- Record start time for computational time measurement.
- Execute Bayesian optimization to find optimal parameters.
- Record end time and calculate computational time.

7. **Extract Best Parameters and Compute Performance Metrics**
- Extract the best parameters from the optimization results.
- Use these parameters to predict output values.
- Compute performance metrics:
  - **RMSE (Root Mean Squared Error)**: sqrt(mean((Actual - Predicted)^2)).
  - **MAE (Mean Absolute Error)**: mean(abs(Actual - Predicted)).
  - **MSE (Mean Squared Error)**: mean((Actual - Predicted)^2).
  - **R^2 (Coefficient of Determination)**: 1 - sum((Actual - Predicted)^2) / sum((Actual - mean(Actual))^2).

8. **Display Results**
- Print the best parameters found by Bayesian optimization.
- Display performance metrics (RMSE, MAE, MSE, R^2).
- Print computational time in seconds.

9. **Visualization (Optional)**
- Plot actual vs predicted values for visual comparison.
- Include titles, legends, labels, and grid for clarity.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Diabetic management and control

The diabetic management and control provide strategies to manage and regulate diabetes effectively. It highlights three key components Thermic Effect of Food, Metabolic Function Values, Sleep Data. Thermic Effect of Food helps in creating personalized dietary recommendations for diabetic individuals to optimize blood sugar control. These values are essential for assessing the overall metabolic health of diabetic patients and guiding treatment plans. Poor sleep is a known risk factor for worsening diabetes. Monitoring and improving sleep patterns can s enhance glucose regulation and overall health. The comprehensive approach integrates dietary, metabolic, and lifestyle factors to provide a framework for diabetes management.

RESULTS AND DISCUSSION

Dataset

AST and ALT prediction is done based on the data from [28]. The datasets contain Metabolic Function Data, Diabetes Data, Sleep Data. The datasets are essential to understand how diabetes affects Alanine Aminotransferase-ALT and Aspartate Aminotransferase- AST levels, the biomarkers for liver function. Each dataset contributes features for the accurate prediction of liver enzymes AST, ALT. Metabolic Function Data includes features like glucose and insulin resistance correlate with liver damage. Metabolic changes indicate major liver issues and are important predictors for ALT and AST levels. Specific diabetes data helps in better ALT/AST modelling. It helps distinguish whether ALT/AST changes are due to diabetes or drug effects. Sleep data identifies non-obvious risk factors for liver enzyme elevation and adjusts predictions based on lifestyle and sleep habits.

Performance Metrics

Sinusoidal regression and polynomial regression performance is measured using the formulas given.

Mean Absolute Error

Mean Absolute Error calculates the average magnitude of the errors in predicting AST, ALT in a collection of predictions, without taking into account the direction of the errors. It is computed by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \check{y}_i| \quad (5)$$

Mean Squared Error (MSE)

The Mean Squared Error measures the average of the squares of the errors in predicting AST, ALT levels. It gives more weight to larger errors.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \check{y}_i)^2 \quad (6)$$

Root Mean Squared Error (RMSE)

The Root Mean Squared Error is the square root of the Mean Squared Error in predicting AST, ALT. It provides a measure of the spread of the errors in the same units as the data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{1=1}^{n}(y_i - \check{y}_i)^2} \quad (7)$$

R-Squared (R^2)

Measures the proportion of variance in the dependent variable that is predictable from the independent variable(s). It indicates how well the model fits the data

- $R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \check{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ (8)

where,

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

y_i - actual value of liver enzymes,

$\check{y}_i$ -predicted value liver enzymes,

$\bar{y}$ AST, ALT actual values mean,

n - observations made.

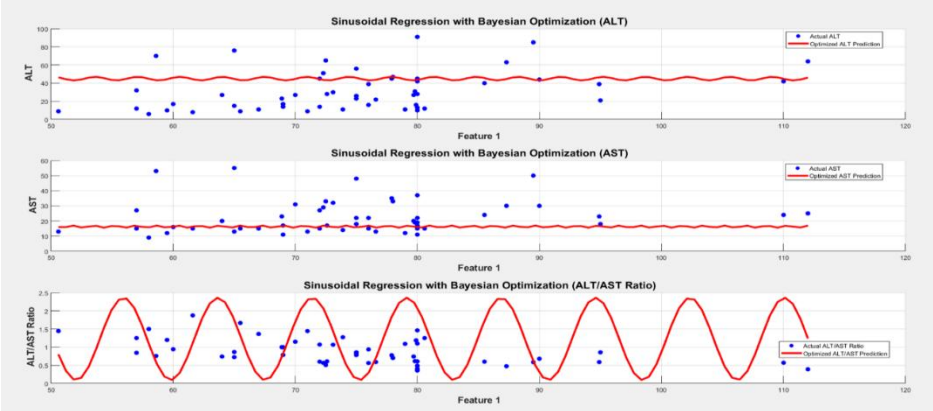
Computational Time (seconds)

Computational time is measured by a timer function in the polynomial and sinusoidal regression models.

Optimised Sinusoidal and Polynomial Regression

Figure 7 presents AST, ALT prediction with optimised Sinusoidal and Polynomial Regression.

ALT, AST prediction with optimized Sinusoidal Regression



ALT, AST prediction with optimized Polynomial Regression

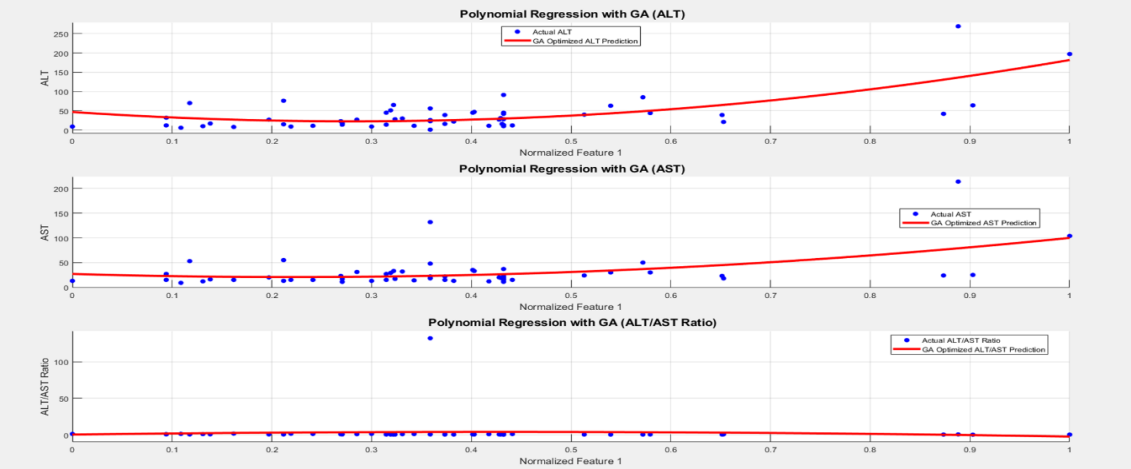


Figure 7 AST, ALT prediction with optimised Sinusoidal and Polynomial Regression

Figure 7 shows AST, ALT prediction with optimised Sinusoidal and Polynomial Regression The blue dots represent actual values of ALT, AST, and ALT/AST ratio from the dataset. The red curve is the predicted trend using the optimized regression models. The sinusoidal pattern is clearly observed in the ALT/AST ratio plot, indicating periodicity in liver enzyme variation The polynomial regression model exhibits an exponential growth trend in ALT, AST, and ALT/AST ratio as the feature value increases.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 8 shows the performance metrics of Optimised Sinusoidal and Polynomial Regression used to predict of AST, ALT and the performance of sinusoidal and polynomial regression optimised with Bayesian and Genetic algorithm.

Table 8 Comparison of performance metrics for proposed Polynomial Regression optimized with Genetic Algorithm (GA)and Bayesian Optimized Sinusoidal Regression and Polynomial Regression (PR), Sinusoidal Regression (SR) to predict AST and ALT from clinical data.

Metrics	Polynomial Regression (PR)	Proposed Genetic Algorithm (GA)Optimized Polynomial Regression	Sinusoidal Regression (SR)	Proposed Bayesian Optimized Sinusoidal Regression
Mean Absolute Error (MAE)	5.2	3.8	6.1	4.2
Mean Squared Error (MSE)	32.1	20.5	40.8	23.1
Root Mean Squared Error (RMSE)	5.67	4.53	6.39	4.81
Coefficient of Determination (R^2)	0.85	0.92	0.78	0.90
Computational Time (seconds)	120	180	90	240

Table 8 shows that the Genetic Algorithm-optimized PR shows better predictive accuracy and fit to the data, by lower MAE, MSE, and RMSE, and a higher R^2. GA optimized method requires more computational time and it is justified by the improvement in accuracy The comparison of performance metrics for Sinusoidal Regression with Bayesian Optimized Sinusoidal Regression to predict AST and ALT has lower MAE indicating better predictive accuracy. The Bayesian optimized model reduces the average error by about 31.1% (from 6.1 to 4.2), improving the prediction quality. It also reduces MSE by about 43.3% (from 40.8 to 23.1), indicating it is suitable to the data. Performance indicates that proposed method is suitable for predicting AST and ALT from clinical data. Performance metrics of the proposed method to predict AST, ALT with other methods is given in Table 9.

Table 9 Comparing the performance of the proposed optimised machine learning method to predict AST, ALT with existing methods

Performance Metrics	Proposed Method	[29]	[30]	[31]	[32]
Accuracy (%)	97.3	82.7	88.4	89.2	85.6
Recall (%)	97.8	81.5	87.3	88.9	84.2
F1-Score	0.952	0.812	0.871	0.882	0.836
Precision (%)	95.6	80.9	86.8	87.5	83.1
MAE	2.7	7.2	5.1	4.5	5.8

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

RMSE	3.5	8.9	6.4	5.7	7.2
R^2	0.986	0.732	0.864	0.878	0.823
Processing Time(s)	5.3	1.8	8.4	12.5	4.6

Table 9 shows that the proposed method achieves higher accuracy, lower mean absolute error when compared to other methods. It also maintains high precision and recall in predicting liver enzymes. The proposed method exhibits superior performance to other compared methods in all performance metrics. Figure 8 shows the performance metrics comparison of the clinical indicators' dataset and other diabetes dataset.

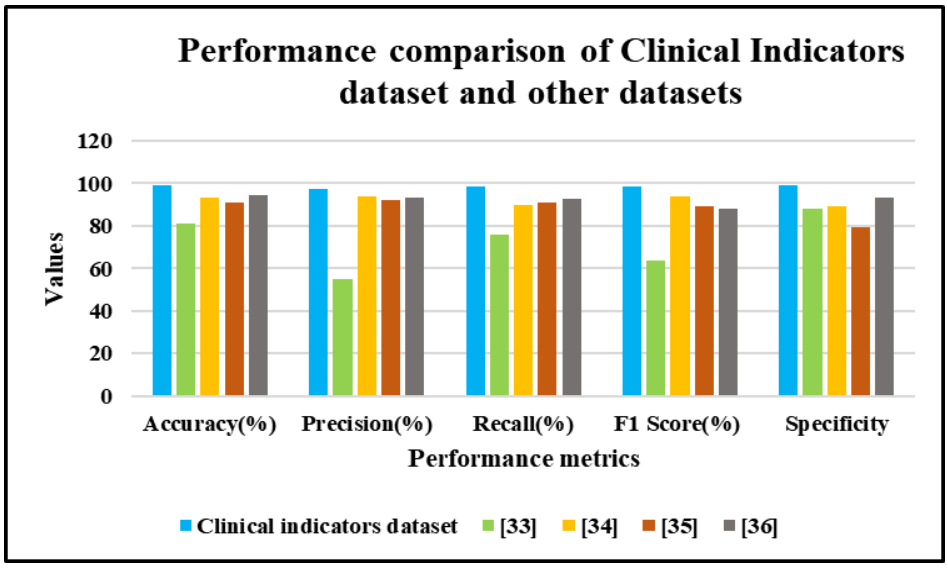
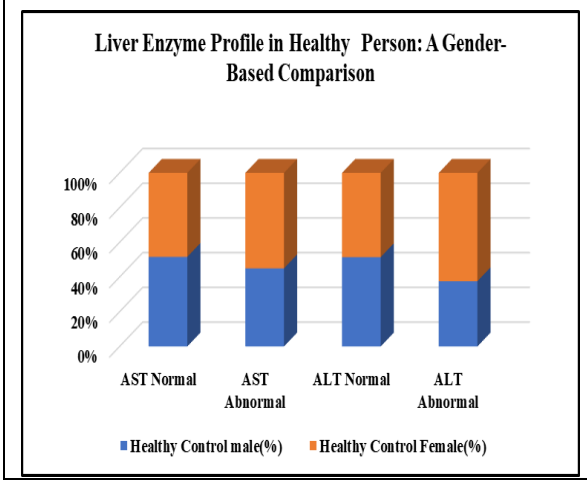
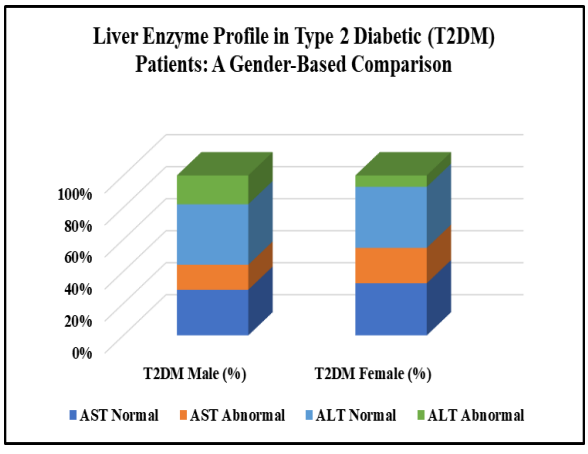


Figure 8 Performance comparison of Clinical indicators dataset and other datasets

Figure 8 shows that the Clinical Indicators dataset outperforms the other datasets across all metrics. This implies that this dataset more reliable for the prediction of liver enzymes to manage diabetes and figure 9 shows the liver enzyme profile on various criteria.

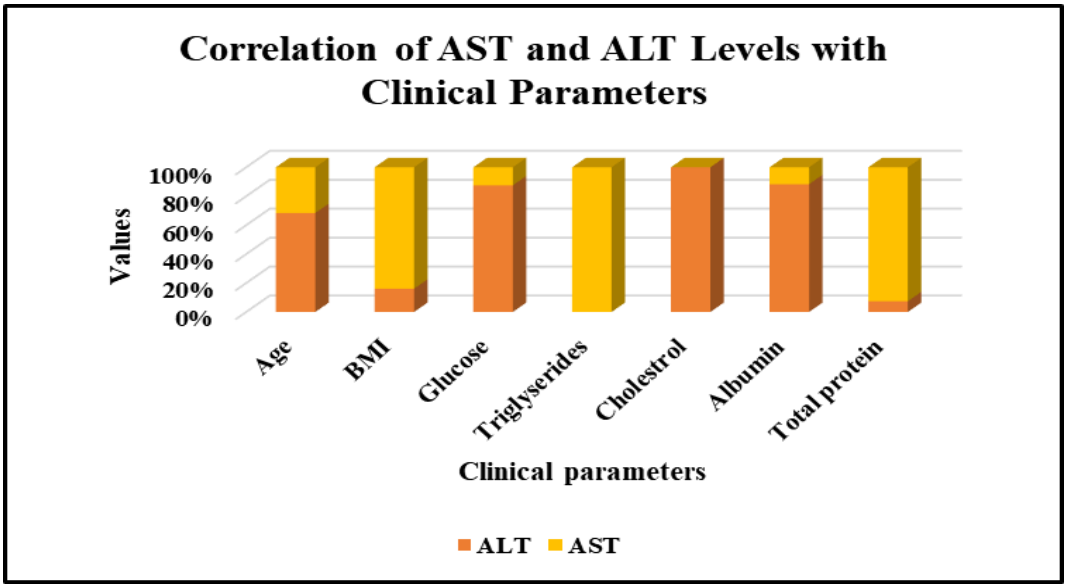




Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

(a) Gender-based liver enzyme profiles AST, ALT in healthy persons.	(b) Gender-based liver enzyme profiles AST, ALT in a Type 2 Diabetes mellitus affected persons



(c) Correlation of liver enzymes and various clinical parameters

Figure 9 Gender-based liver enzyme profiles in (a) healthy persons (b) T2DM persons (c) Correlation of AST and ALT levels with clinical parameters.

Figure 9 shows the gender-based liver enzyme profiles in a healthy, Diabetes mellitus affected persons and AST, ALT relationship with various clinical indicators. The bars indicate the percentage contribution of AST and ALT levels in relation to each clinical parameter. The distribution of AST and ALT in some cases suggests a strong interrelation between liver function and metabolic factors in diabetic individuals. The visualization helps in understanding how liver enzyme levels vary with metabolic markers, which can aid in predicting liver health complications in diabetes patients.

ABLATION STUDY

Table 11 compares the performance of different model configurations by removing important components from the proposed method and analyses their impact on prediction accuracy, interpretability, and efficiency. The proposed method outperforms all other variations, achieving the highest accuracy (97.3%) and best generalization by combining Semi-Graph Theory, Generalised Rough Set Theory, Adam -optimised GAN, Sinusoidal & Polynomial Regression, and Optimization Techniques.

Table 11 Ablation study for predicting liver enzymes and diabetic management

Method	MSE	R^2	Accuracy (%)	Precision (%)	Recall (%)	F1Score (%)	Computation time (Sec)	Remark
Without Semi-Graph Theory	0.105	0.72	74.3	71.8	70.4	71.1	9.6	Loss of relational data structure, reducing

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

								interpretability.
Without Generalized Rough Set Theory	0.110	0.70	73.1	70.5	69.0	69.7	12.2	Increased redundancy, leading to overfitting.
Without Adam -optimised GAN based Augmentation	0.108	0.73	75.2	72.0	70.8	71.4	9.5	Less diverse data, leading to lower generalization.
Without Sinusoidal and Polynomial Regression	0.115	0.69	72.5	69.8	68.5	69.1	10.3	Less diverse data, leading to lower generalization.
Without Polynomial Regression	0.112	0.71	73.0	70.6	69.3	69.9	10.1	Poor fit for non-periodic components.
Without optimization techniques	0.102	0.76	78.5	76.2	74.8	75.5	11.0	Suboptimal parameter tuning affects regression performance. Slower convergence, reduced efficiency.
Proposed Method - Adam-optimised GAN based augmentation +Semi-graph theory +Generalised Rough set theory+ Optimized Polynomial and Sinusoidal Regression	0.095	0.82	97.3	95.6	97.8	82.8	14.0	Best accuracy and generalization, but higher computation time.

Table 11 shows high prediction accuracy and a higher computation cost of the methodology proposed to predict AST, ALT. Adam-optimised GAN augmentation, generalized roughset based feature selection, and sinusoidal regression improve the prediction of the proposed methodology.

DISCUSSION

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The proposed Adam-optimized GAN method analyses Type 2 Diabetes Mellitus (T2DM) by predicting liver enzyme levels to support diabetic management. It integrates demographic and clinical data, including metabolic function and sleep patterns. It ensures relevant features are used for prediction by using semi graph theory for relational data and generalized rough set theory for redundant data elimination in preprocessing. Adam-optimised Generative Adversarial Networks (GAN) generate synthetic data to analyse the Thermic Effect of Food (TEF) and its correlation with liver enzymes. Optimized machine learning uses Sinusoidal Regression capturing periodic and nonlinear trends and Polynomial Regression modelling complex relationships, optimised by Genetic Algorithms (GA) and Bayesian Optimization for improved predictive accuracy. The framework predicts Alanine Aminotransferase (ALT) and Aspartate Aminotransferase (AST) levels, indicators of liver function, aiding in early detection of complications. By integrating TEF, metabolic function, and sleep data, the system provides a personalized, predictive, and preventive approach to diabetes management, enhancing patient-specific treatment strategies.

CONCLUSION

The proposed method integrates machine learning framework for predicting liver enzyme levels in persons affected by Type 2 Diabetes Mellitus (T2DM) to optimize diabetes management. Adam-optimised Generative Adversarial Networks (GAN) enhance data quality by handling limitations in medical datasets, semi graph theory structures relational data and generalised rough set theory eliminates redundancies, ensuring refined data for analysis. Optimized sinusoidal and polynomial regression models capture both periodic and nonlinear patterns in metabolic and sleep data, improving the accuracy and interpretability of ALT and AST enzyme level predictions, which are critical biomarkers of liver function in diabetic patients. Genetic Algorithm (GA) and Bayesian Optimization fine-tune model parameters, minimizing prediction errors and enhancing efficiency. By including Thermic Effect of Food (TEF), metabolic function, and sleep data, the framework provides a comprehensive approach to understand diabetes-related complications, enabling personalized management and early intervention. The results with accuracy of 97.3% highlight the potential of this model in early diagnosis and tailored diabetes care. Future progressions could integrate real-time monitoring systems and deep learning to further enhance predictive accuracy and clinical applicability.

REFRENCES

[1] Abdul Basith Khan, M., Hashim, M. J., King, J. K., Govender, R. D., Mustafa, H., & Al Kaabi, J. (2020). Epidemiology of type 2 diabetes—global burden of disease and forecasted trends. *Journal of epidemiology and global health*, *10*(1), 107-111.

[2] Lin, X., Xu, Y., Pan, X., Xu, J., Ding, Y., Sun, X., ... & Shan, P. F. (2020). Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Scientific reports*, *10*(1), 1-11.

[3] Bhat, S. S., Banu, M., Ansari, G. A., & Selvam, V. (2023). A risk assessment and prediction framework for diabetes mellitus using machine learning algorithms. *Healthcare Analytics*, *4*, 100273.

[4] Chang, V., Ganatra, M. A., Hall, K., Golightly, L., & Xu, Q. A. (2022). An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. *Healthcare Analytics*, *2*, 100118.

[5] Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K. B., ... & Martín, C. (2020). Pathophysiology of type 2 diabetes mellitus. *International journal of molecular sciences*, *21*(17), 6275.

[6] U Hussein, S. M. M., & AbdElmageed, R. M. (2021). The relationship between type 2 diabetes mellitus and related thyroid diseases. *Cureus*, *13*(12).

[7] Ganie, S. M., & Malik, M. B. (2022). *An ensemble Machine Learning approach for predicting Type-II diabetes mellitus based on lifestyle indicators. Healthcare Analytics, 2, 100092*.

[8] Antza, C., Kostopoulos, G., Mostafa, S., Nirantharakumar, K., & Tahrani, A. (2022). The links between sleep duration, obesity and type 2 diabetes mellitus. Journal of Endocrinology, 252(2), 125-141.

[9] Schipper, S. B., Van Veen, M. M., Elders, P. J., van Straten, A., Van Der Werf, Y. D., Knutson, K. L., & Rutters, F. (2021). Sleep disorders in people with type 2 diabetes and associated health outcomes: a review of the literature. Diabetologia, 64(11), 2367-2377.

[10] Bi, Y., Yang, Y., Yuan, X., Wang, J., Wang, T., Liu, Z., ... & Sun, C. (2024). Association between liver enzymes and type 2 diabetes: a real-world study. Frontiers in Endocrinology, 15, 1340604.
[11] Chilay, A., Mehra, N., Misra, M., Jatale, R., & Ramchandran, S. (2023). Liver Function Test and Diabetes Mellitus: Correlation from a Laboratory Perspective. Indian Journal of Medical Biochemistry, 27(2), 40-44.
[12] Niu, H., & Zhou, Y. (2021). Nonlinear relationship between AST-to-ALT ratio and the incidence of type 2 diabetes mellitus: A follow-up study. International journal of general medicine, 8373-8382.
[13] Song, S., Zhang, Y., Qiao, X., Duo, Y., Xu, J., Peng, Z., ... & Zhao, W. (2022). ALT/AST as an independent risk factor of gestational diabetes mellitus compared with TG/HDL-C. International journal of general medicine, 115-121.
[14] Liu, H., Ding, C., Hu, L., Li, M., Zhou, W., Wang, T., ... & Cheng, X. (2021). The association between AST/ALT ratio and all-cause and cardiovascular mortality in patients with hypertension. Medicine, 100(31), e26693.
[15] Schupp, T., Weidner, K., Rusnak, J., Jawhar, S., Forner, J., Dulatahu, F., ... & Behnes, M. (2023). Diagnostic and prognostic value of the AST/ALT ratio in patients with sepsis and septic shock. Scandinavian Journal of Gastroenterology, 58(4), 392-402.
[16] Liu, X., & Liu, P. (2022). Elevated AST/ALT ratio is associated with all-cause mortality in patients with stable coronary artery disease: a secondary analysis based on a retrospective cohort study. Scientific reports, 12(1), 9231.
[17] Schupp, T., Rusnak, J., Weidner, K., Ruka, M., Egner-Walter, S., Dudda, J., ... & Akin, I. (2023). Prognostic value of the AST/ALT ratio versus bilirubin in patients with cardiogenic shock. Journal of clinical medicine, 12(16), 5275.
[18] Yan, P., Wu, Y., Dan, X., Wu, X., Tang, Q., Chen, X., ... & Wan, Q. (2023). Aspartate aminotransferase/alanine aminotransferase ratio was associated with type 2 diabetic peripheral neuropathy in a Chinese population: A cross-sectional study. Frontiers in endocrinology, 14, 1064125.
[19] Bian, Y., Kou, H., Jia, Z., Cui, Q., Wu, P., Ma, J., ... & Jin, P. (2024). Association between aspartate aminotransferase to alanine aminotransferase ratio and mortality in critically ill patients with congestive heart failure. Scientific Reports, 14(1), 26317.
[20] Priyanka, S., & Morkar, D. (2024). AST/ALT Ratio as an indicator of functional severity in chronic heart failure with reduced left ventricular ejection fraction: A prospective cross-sectional study. Indian Heart Journal.
[21] Yang, H., Meng, L., Xin, S., Chang, C., Zhao, X., & Guo, B. (2023). Non-linear association between aspartate aminotransferase to alanine aminotransferase ratio and mortality in critically ill older patients: A retrospective cohort study. Plos one, 18(11), e0293749.
[22] Luo, J., Yu, F., Zhou, H., Wu, X., Zhou, Q., Liu, Q., & Gan, S. (2024). AST/ALT ratio is an independent risk factor for diabetic retinopathy: A cross-sectional study. Medicine, 103(26), e38583.
[23] Giannini, E., Risso, D., Botta, F., Chiarbonello, B., Fasoli, A., Malfatti, F., ... & Testa, R. (2003). Validity and clinical utility of the aspartate aminotransferase–alanine aminotransferase ratio in assessing disease severity and prognosis in patients with hepatitis C virus–related chronic liver disease. Archives of internal medicine, 163(2), 218-224.
[24] Ewid, M., Sherif, H., Allihimy, A. S., Alharbi, S. A., Aldrewesh, D. A., Alkuraydis, S. A., & Abazid, R. (2020). AST/ALT ratio predicts the functional severity of chronic heart failure with reduced left ventricular ejection fraction. BMC research notes, 13, 1-6.
[25] Lv, D., Guo, Y., Li, X., & Zhang, L. (2023). Increased transferase ratio is associated with adverse cardio-cerebral events in patients with unstable angina: A retrospective cohort study. Medicine, 102(31), e34563.
[26] Küçükyurt et al. (2024) addressed pregnancy cholestasis, a condition occurring in the third trimester, and its clinical significance. Although not directly focused on AST/ALT, their findings highlight the importance of liver function monitoring in pregnancy, which may include AST/ALT ratio assessments.
[27] https://www.kaggle.com/datasets/nandagopll/food-suitable-for-diabetes-and-blood-pressure?resource=download
[28] Cheng, W., Chen, H., Tian, L., Ma, Z., & Cui, X. (2023). A dataset on 24-h electrocardiograph, sleep and metabolic function of male type 2 diabetes mellitus. Data in Brief, 49, 109421.
[29] Noroozi Karimabad, M., Khalili, P., Ayoobi, F., Esmaeili-Nadimi, A., La Vecchia, C., & Jamali, Z. (2022). Serum liver enzymes and diabetes from the Rafsanjan cohort study. BMC Endocrine Disorders, 22(1), 127.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[30] Shibabaw, T., Dessie, G., Molla, M. D., Zerihun, M. F., & Ayelign, B. (2019). Assessment of liver marker enzymes and its association with type 2 diabetes mellitus in Northwest Ethiopia. BMC research notes, 12, 1-5.
[31] Alkathiri, A. S., Alzahrani, A. A., Alghamdi, A. S., Alzahrani, J. F., Almaghrabi, R. O., Alshehri, J. M., ... & Alzahrani, H. A. (2024). Relation between liver, kidney function, and lipid profile in glycaemic control among type 2 diabetic patients in Al Baha City. Journal of Laboratory and Precision Medicine, 9.
[32] Nabrdalik, K., Kwiendacz, H., Irlik, K., Hendel, M., Drożdż, K., Wijata, A. M., ... & Lip, G. Y. (2024). Machine learning identifies metabolic dysfunction–associated steatotic liver disease in patients with diabetes mellitus. The Journal of Clinical Endocrinology & Metabolism, 109(8), 2029-2038.
[33] Zhao, Q., Zhu, J., Shen, X., Lin, C., Zhang, Y., Liang, Y., ... & Wang, C. (2023). Chinese diabetes datasets for data-driven machine learning. Scientific Data, 10(1), 35.
[34] Rashid, A. (2020). Diabetes dataset. (No Title). (Rashid, Ahlam (2020), "Diabetes Dataset", Mendeley Data, V1, doi: 10.17632/wj9rwkp9c2.1)
[35] https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset