JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Risk Factors Analysis and Predictive Modelling of Diabetic Neuropathy Using Machine Learning Algorithms

[1]Aminu Abdulmajeed Mikailu, [2]Muhammad Garba, [3]Anas Tukur Balarabe, [4]Anas Ahmed Sabir
[1]*Alqalam Islamic University, Katsina*
[2]*Kebbi State University of Science and Technology, Aliero*
[3]*Sokoto State University, Sokoto*
[4]*Usmanu Danfodiyo University Teaching Hospital Sokoto*

ABSTRACT

Diabetic neuropathy is a common and severe complication of diabetes mellitus, requiring early detection for effective management. This study investigated the predictive potential of machine learning algorithms—Naive Bayes, Support Vector Machine (SVM), and Decision Tree (J48)—using clinical and demographic data from 356 patients at Usmanu Danfodiyo University Teaching Hospital, Sokoto, Nigeria. The Decision Tree (J48) model demonstrated superior performance with an accuracy of 95.06%, followed by SVM (91.56%) and Naive Bayes (83.29%). Key predictive factors included poor glycemic control, duration of diabetes, family history, cardiovascular conditions, and neurological symptoms such as burning pain and muscle weakness. The findings highlight the utility of machine learning in supporting early diagnosis and individualized care. Incorporating these models into clinical practice could significantly enhance the management of diabetic neuropathy, reduce long-term complications, and optimize healthcare resources, particularly in low-resource settings.

ARTICLE INFO
Article History
Received: September, 2025
Received in revised form: October, 2025
Accepted: December, 2025
Published online: January, 2026

KEYWORDS
Diabetic Neuropathy; Machine Learning Algorithms Machine Learning; Predictive Modelling; Risk Factor Analysis

INTRODUCTION

Diabetes mellitus (DM) is a chronic, multifactorial metabolic disorder marked by deficiencies in insulin secretion, insulin action, or both. Insulin, secreted by the pancreas, plays a central role in glucose metabolism by facilitating the uptake of glucose from the bloodstream into body cells for energy production. Impaired insulin function leads to elevated blood glucose levels, which, if left uncontrolled, can result in a range of systemic complications affecting vital organs such as the eyes, kidneys, heart, and peripheral nerves. The burden of diabetes is especially profound in low- and middle-income countries, where limited healthcare access exacerbates its long-term complications (Sanyaolu et al., 2023).

Globally, the prevalence of diabetes has risen at an alarming rate over the past three decades, increasing nearly fourfold and now ranking as the ninth leading cause of death. Recent estimates indicate that approximately one in eleven adults is currently living with diabetes, with type 2 diabetes accounting for 90% of all cases. In 2019 alone, diabetes was directly responsible for 1.5 million deaths, with nearly half of these occurring in individuals under the age of 70. Additionally, hyperglycemia contributed significantly to deaths from cardiovascular and kidney-related diseases (World Health Organization, 2023). The highest burdens are concentrated in Asia, particularly in China and India (Ismail & Ali, 2022).

Modifiable lifestyle factors—such as poor diet, physical inactivity, obesity, and harmful health behaviors—are major contributors to the onset and progression of diabetes. Strategies including regular exercise, balanced nutrition, and weight management have been shown to reduce diabetes risk and severity (Suryasa, Rodriguez-Gamez, & Koldoris, 2021). The three primary

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

forms of diabetes include: type 1 diabetes (autoimmune destruction of insulin-producing cells), type 2 diabetes (insulin resistance or deficiency), and gestational diabetes (hyperglycemia during pregnancy).

Among the most debilitating complications of diabetes is diabetic neuropathy (DN)—a chronic condition resulting from damage to the peripheral nerves due to prolonged hyperglycemia. DN is a significant cause of morbidity and disability worldwide and is considered the most common form of neuropathy in diabetic patients, particularly in high-income countries. Studies have reported DN prevalence rates ranging from 6% to 51% in North America and Europe, and up to 46% in Africa (Azeez et al., 2022). The most frequent manifestation is distal symmetric polyneuropathy, often presenting with sensory loss, dysesthesia, or pain in the extremities. Risk factors include poor glycemic control, long diabetes duration, aging, hypertension, dyslipidemia, and associated microvascular complications such as diabetic retinopathy.

Diagnostic approaches for DN include nerve conduction studies, which are considered the gold standard, along with physical examinations, scoring instruments like the Michigan Neuropathy Screening Instrument, vibration perception tests, and monofilament assessments. However, these diagnostic tools are often expensive, time-consuming, or reliant on subjective interpretation, which limits their utility in resource-constrained settings (Bondar et al., 2021).

In recent years, the digitization of healthcare systems has resulted in the generation of large volumes of clinical data. This data boom has spurred the integration of data science techniques—particularly machine learning (ML)—into medical research and practice. ML techniques have demonstrated immense potential in predicting the onset of chronic diseases, optimizing treatment pathways, and improving diagnostic accuracy. Applications in healthcare include predictive modeling, clinical decision support, fraud detection, patient stratification, and outcome forecasting (Jain & Devendran, 2021; Sisodia & Sisodia, 2018).

Despite the growing body of research leveraging ML for diabetes-related conditions—including tropical diabetic hand syndrome, imbalanced data classification, and glycemic risk prediction—relatively few studies have focused on the early detection of diabetic neuropathy using machine learning. Most existing models either overlook neuropathy-specific features or rely on post-complication data, limiting their effectiveness in real-time clinical settings. Furthermore, traditional approaches are hampered by the lack of integration between electronic health records and predictive algorithms (Wang et al., 2019; Ahmed et al., 2022).

Given the severity and growing prevalence of diabetic neuropathy, coupled with limitations in early diagnostic tools, there is a pressing need for innovative, data-driven approaches that can accurately predict the onset of DN. This study addresses this gap by developing a machine learning-based predictive model using algorithms such as Decision Tree, Naive Bayes, and Support Vector Machine (SVM). These models are applied to clinical and demographic data obtained from diabetic patients at Usmanu Danfodiyo University Teaching Hospital in Sokoto, Nigeria.

The objectives of this research are fourfold: (1) to identify key risk factors associated with diabetic neuropathy; (2) to evaluate the influence of these factors on neuropathy progression; (3) to determine the local prevalence of diabetic neuropathy within the study population; and (4) to develop a clinical algorithm capable of accurately predicting DN onset. By achieving these aims, the study seeks to improve early diagnosis, support timely clinical intervention, and ultimately reduce the burden of diabetic complications in resource-limited settings.

LITERATURE REVIEW

Diabetes mellitus (DM) is a chronic metabolic disorder that has emerged as a significant global health concern, with its prevalence rising sharply across both developed and developing countries. Sanyaolu et al. (2023)

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

highlighted the disproportionate burden of diabetes in low- and middle-income nations, attributing increased disability and mortality rates to the disease's long-term complications. Their study emphasized the importance of early intervention strategies, including healthy diets, regular physical activity, and the use of pharmacological treatments. Similarly, Ismail and Ali (2022) noted a global surge in diabetes incidence, reporting 1.5 million diabetes-related deaths in 2012 alone, nearly half of which occurred before the age of 70. These findings support the need for individualized glycemic control plans that are responsive to patient comorbidities and life expectancy.

Suryasa et al. (2021) and Olokoba et al. (2012) reinforced the role of rapid urbanization, aging, and lifestyle changes as key drivers of the growing diabetes epidemic. They projected that more than 366 million individuals would be living with diabetes by 2030, with developing nations bearing the brunt of the increase. Chaudhary and Tyagi (2018) contributed further by exploring the pathophysiology of diabetes, linking the disease to obesity, hormonal dysregulation, and genetic mutations. Their research detailed common symptoms—including fatigue, weight loss, and peripheral numbness—which often precede more severe complications such as diabetic neuropathy (DN).

Diabetic neuropathy is widely acknowledged as one of the most prevalent and debilitating complications associated with diabetes. Nascimento et al. (2016) identified distal symmetric polyneuropathy (DSP) as the most common subtype of DN, affecting up to 50% of diabetic patients. Their comprehensive review discussed the roles of metabolic dysregulation, microvascular dysfunction, and inflammatory responses in the progression of DN. Diagnostic approaches such as nerve conduction studies and skin biopsies were emphasized, alongside treatment strategies involving strict glycemic control, antioxidant supplementation like α-lipoic acid, and symptom management using antidepressants and anticonvulsants.

Expanding on the molecular understanding of DN, Zhu et al. (2024) examined the involvement of oxidative stress, insulin resistance, and intracellular signaling pathways—such as mTOR and Wnt/β-catenin—in diabetic peripheral neuropathy (DPN). Their findings exposed a persistent gap in the availability of personalized treatment options, especially concerning the differing pathogenesis between Type 1 and Type 2 diabetes. These insights highlight the necessity for more nuanced therapeutic approaches.

In recent years, machine learning (ML) techniques have been increasingly applied to improve the diagnosis and risk prediction of diabetic neuropathy. Xu et al. (2020) developed ML-based predictive models for DN in Type 1 diabetic patients, utilizing clinical variables such as age, HbA1c, and disease duration. While the models demonstrated promising performance, the authors noted a lack of external validation and limited dataset diversity. Dubey et al. (2020) introduced an artificial neural network (ANN) model that exhibited high predictive accuracy for DN risk classification. Similarly, Baskozos et al. (2022) employed Random Forest and Adaptive Regression Splines to effectively differentiate between painful and painless forms of DN, showcasing the potential of ML in clinical stratification.

Additional technological advances have focused on non-invasive diagnostic tools. Carmichael et al. (2021) reviewed the application of corneal confocal microscopy (CCM) and intra-epidermal nerve fiber density (IENFD) as early biomarkers for DN. Although these techniques offer diagnostic sensitivity, concerns over cost, accessibility, and standardization have limited their widespread adoption in routine clinical settings. In a related context, Tiwari et al. (2008) explored the use of ensemble machine learning models—including Random Forest, Naïve Bayes, and AdaBoost—for predicting Tropical Diabetic Hand Syndrome (TDHS), a rare but severe diabetes-related complication. Their study demonstrated the superior performance of ensemble approaches compared to single-model algorithms, reinforcing the value of computational techniques in medical risk prediction.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

A growing body of empirical evidence has established that diabetic neuropathy is influenced by a combination of non-modifiable and modifiable risk factors. The duration of diabetes remains a strong predictor, with studies by Aliyu et al. (2023), Azeez et al. (2022), and Bondar et al. (2021) confirming that prolonged hyperglycemia exacerbates nerve damage. Aging also increases susceptibility, as metabolic stress and impaired neuronal repair mechanisms worsen with time, as noted by Baldimtsi et al. (2024). Gender- and ethnicity-related differences have also been documented; Akinci et al. (2020) and the TODAY Study Group (2022) found that hormonal and environmental variables can influence disease expression and progression across different demographic groups.

Among modifiable factors, poor glycemic control remains paramount. Chronic hyperglycemia accelerates oxidative stress and inflammatory processes, leading to progressive nerve degeneration (Bondar et al., 2021; Ponirakis et al., 2022). Obesity, hypertension, and dyslipidemia have been shown to compound DN risk by promoting insulin resistance, vascular damage, and chronic inflammation (Azeez et al., 2022; Liu et al., 2022; Pastore et al., 2022). Lifestyle behaviors, including smoking, alcohol use, and physical inactivity, further contribute to microvascular dysfunction and exacerbate neuropathic complications (Aliyu et al., 2023; Bondar et al., 2021).

Recent literature has also drawn attention to emerging biological factors. Systemic inflammation and elevated oxidative stress markers, such as inflammatory cytokines and reactive oxygen species (ROS), have been identified in DN patients (Liu et al., 2022). Moreover, microvascular complications, including diabetic nephropathy and retinopathy, share pathogenic mechanisms with DN, increasing the likelihood of concurrent conditions (Galiero et al., 2023; Pastore et al., 2022). Genetic predispositions and autoimmune responses are also being explored as predictors of DN, with studies suggesting links between specific genetic markers and susceptibility to nerve damage (Bondar et al., 2021; Galiero et al., 2023).

Collectively, the empirical evidence positions diabetic neuropathy as a multifactorial condition shaped by a dynamic interplay of metabolic, behavioral, demographic, and genetic variables. While conventional diagnostic approaches remain vital, the integration of machine learning models holds considerable promise for enhancing early detection and personalized care. Nonetheless, the effectiveness of these computational tools depends heavily on data quality, external validation, and clinical integration. The insights gleaned from this body of literature form the foundation for the present study, which aims to address gaps in early DN prediction by developing a machine learning–based model tailored to clinical data from a Nigerian tertiary healthcare setting.

Recent empirical studies have demonstrated the significant potential of data mining and machine learning (ML) techniques in advancing healthcare delivery, particularly in the diagnosis and prediction of diabetes and its complications. Sadiku et al. (2018) described data mining as a transformative process that extracts meaningful patterns from data warehouses to support clinical decision-making. Their findings indicated that data mining has become instrumental in critical areas such as patient care, predictive analytics, cost reduction, and healthcare optimization. The study outlined key phases of the data mining process, including data pre-processing, feature selection, visualization, and outcome integration into clinical workflows.

The application of ML techniques in diabetes prediction has been explored extensively. Bhaskar et al. (2023) developed a Correlational Neural Network (CORNN) model that utilized breath acetone levels—a biomarker correlated with blood glucose—to predict Type 2 diabetes. The model achieved an impressive accuracy of 98.02%, surpassing conventional classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), and Random Forest (RF). In a related study, Khaleel and Al-Bakry (2023) applied ML algorithms to the Pima Indian Diabetes Dataset (PIDD), achieving prediction accuracies of 74%, 77%, and 69% for logistic

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

regression, Naïve Bayes, and KNN, respectively, thus demonstrating variability in algorithmic performance.

Bhat et al. (2022) compared six ML algorithms—Random Forest, Multilayer Perceptron, SVM, Gradient Boost, Decision Tree, and Logistic Regression—for early diabetes prediction. Random Forest yielded the highest predictive accuracy, reaffirming its robustness in classification tasks. Similarly, Jijo and Abdulazeez (2021) reviewed the performance of decision tree algorithms over a three-year period and found that Decision Trees, while interpretable and easy to implement, could attain up to 99.93% accuracy under ideal conditions, albeit with limitations such as overfitting and model instability.

Further performance comparisons were conducted by Ergun and Ilhan (2021), who evaluated eight ML algorithms—including Decision Tree, SVM, RF, XGBoost, KNN, Naïve Bayes, Artificial Neural Networks (ANN), and CNN—on a dataset comprising 250 individuals and 16 clinical features. Their analysis, validated through tenfold cross-validation, revealed CNN as the best performer (97.44% accuracy), while Naïve Bayes showed the least accuracy (88.85%). These findings reinforce the growing preference for deep learning models in complex pattern recognition tasks.

Incorporating semantic technologies, El-Massari et al. (2022) proposed an ontology-enhanced ML approach using the PIDD dataset. Their ontology-based model, developed in Protégé and integrated into the Weka platform, demonstrated superior performance in terms of accuracy, precision, recall, and F-measure compared to six conventional classifiers, highlighting the potential of ontology frameworks to enhance model interpretability and efficiency.

Ahmed et al. (2022) developed a fused ML model combining SVM and ANN to enhance diabetes prediction accuracy. The fusion model achieved a classification accuracy of 94.87%, outperforming its individual components. This supports findings by Ibrahim and Abdulazeez (2021), who reported that ensemble and hybrid learning systems often yield better performance due to their ability to integrate the strengths of multiple algorithms and compensate for individual model limitations.

Regarding the types of ML, Abdullah and Abdulazeez (2021) categorized learning algorithms into supervised, unsupervised, and reinforcement learning. They highlighted SVM as a robust supervised algorithm applicable in fields such as disease diagnosis, face recognition, and intrusion detection, with reported accuracy reaching 97% in certain domains. Additionally, Wang et al. (2019) addressed the challenge of missing and imbalanced data in diabetes prediction using a Naïve Bayes classifier alongside Random Forest. By applying adaptive synthetic sampling and imputation techniques, their model achieved an accuracy of 87.10%, demonstrating the efficacy of preprocessing strategies in enhancing model performance.

Overall, these studies affirm the increasing relevance of machine learning and data mining techniques in modern healthcare. They underscore the importance of algorithm selection, data quality, hybrid systems, and semantic enhancement in building reliable predictive models for diabetes and related conditions. The empirical evidence supports further integration of ML tools into clinical workflows to facilitate early diagnosis, personalized care, and improved patient outcomes.

METHODOLOGY

Study Design

This study employs a quantitative research design to develop a predictive model for diabetic neuropathy using supervised machine learning algorithms—Naïve Bayes, Support Vector Machine (SVM), and Decision Tree. The choice of a quantitative approach is motivated by the measurable nature of risk factors associated with diabetic neuropathy, such as glycated hemoglobin levels, duration of diabetes, age, and blood pressure. The Waikato Environment for Knowledge Analysis (WEKA) was adopted as the primary simulation environment due to its robust suite of machine learning tools, ease of integration, and suitability for clinical classification tasks.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Study Area and Population

The research was conducted at Usmanu Danfodiyo University Teaching Hospital (UDUTH), Sokoto, Nigeria. The hospital, a tertiary healthcare facility and academic center, provides specialized care to a diverse population in northwestern Nigeria. The study population comprised patients diagnosed with diabetes mellitus attending the endocrinology unit. The inclusion criteria covered adult patients (aged ≥18 years) with a confirmed diagnosis of Type 1 or Type 2 diabetes mellitus, irrespective of gender or comorbid conditions.

Sampling and Sample Size Determination

A simple random sampling technique was employed to ensure an unbiased and representative sample. The sample size was calculated using Cochran's formula for large populations, assuming a 95% confidence level, a 5% margin of error, and a population proportion (p) of 0.5. Given an estimated diabetic patient population of 2,600 over a 12-month period, the sample size was adjusted for a finite population:

Cochran's Formula (Infinite Population):

$$n_0 = \frac{Z^2 \bullet p.q}{e^2} = \frac{(1.96)^2 \bullet 0.5 \bullet 0.5}{(0.5)^2} = 384$$

Adjusting for a finite population (N = 2600):

$$n = \frac{n_o}{1 + \frac{n_o - 1}{N}} = \frac{384}{1 + \frac{383}{2600}} \approx 335$$

Thus, the final sample comprised 335 diabetic patients**.**

Data Collection

Primary data were collected via structured questionnaires and semi-structured interviews. The questionnaire contained both closed-ended and Likert-scale items capturing demographic information, diabetes history, glycemic control, symptom severity, lifestyle factors, and medical history relevant to diabetic neuropathy. Interviews offered qualitative insight to supplement the quantitative dataset. Prior ethical clearance and informed consent were obtained from all participants.

Data Preprocessing and Feature Selection

Preprocessing steps included data cleaning to address missing values and remove redundancies. Numerical attributes with missing values were imputed using the mean or median, while categorical attributes were handled using mode imputation. Feature selection techniques, including Pearson's correlation coefficient, Chi-square test, and Recursive Feature Elimination (RFE), were applied to identify the most relevant predictors and reduce dimensionality, thereby improving model interpretability and performance.

Classification Algorithms

This study employed three supervised machine learning algorithms—Naïve Bayes, Support Vector Machine (SVM), and Decision Tree—for the classification of diabetic neuropathy risk. The Naïve Bayes algorithm, a probabilistic classifier based on Bayes' Theorem, assumes feature independence and is particularly effective for high-dimensional datasets and scenarios involving missing values due to its computational efficiency. The SVM algorithm constructs an optimal hyperplane in a high-dimensional space to distinguish between classes, employing kernel functions such as the radial basis function (RBF) to handle non-linearly separable data. The Decision Tree algorithm creates an interpretable, rule-based hierarchical structure by recursively partitioning the feature space, using metrics such as entropy and information gain to guide node splitting and improve classification performance.

Model Evaluation

Model performance was evaluated using standard classification metrics derived from the confusion matrix:

1. **Accuracy**: $$\frac{TP + TN}{TP + TN + FP + FN}$$
2. **Precision**: $\frac{TP}{TP + FP}$

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

3. **Recall (Sensitivity)**: $\frac{TP}{TP + FN}$
4. **F1-Score**: Harmonic mean of precision and recall
5. **False Positive Rate (FPR)**: $\frac{FP}{FP + TN}$
6. **Area Under the ROC Curve (AUC)**: Evaluates the model's ability to distinguish between classes.

All models were trained and validated using 10-fold cross-validation to minimize overfitting and ensure robustness. WEKA and IBM SPSS were used for machine learning implementation and descriptive statistical analysis, respectively.

RESULTS

Table 1 presents summary statistics of the age distribution among the 356 respondents. Participants ranged in age from 30 to 80 years, with a mean age of 60.94 years (SD = 13.10), indicating a broad age range and variability. This suggests the sample comprises a demographically diverse group, enhancing the generalizability of the study findings across various adult age groups.

Table 1: Age Distribution

Age	N	Minimum	Maximum	Mean	Std. Deviation
	356	30	80	60.94	13.096
Valid N (listwise)	356				

As shown in Table 2, the median age was 63, while the mode, or most frequently occurring age, was 77. The slight difference between the mean and median suggests a mild positive skew, with more participants clustering in the older age categories. The high mode underscores the predominance of older adults in the sample, a factor relevant in diabetes-related studies, as the risk for complications like neuropathy tends to increase with age.

Table 2: Age Central Tendency Measures

Valid	356
Missing	0
Mean	60.94
Median	63.00
Mode	77
Valid	356

As presented in Table 3, out of the 356 participants surveyed, 197 (55.3%) were female and 159 (44.7%) were male, indicating a slightly higher representation of females in the study sample. This aligns with common observations in health-related survey studies where women tend to be more represented due to greater health-seeking behavior. The distribution, illustrated in Figure 4.2, supports the demographic diversity of the sample and suggests that gender may be a factor worth exploring in relation to diabetes management and neuropathic outcomes

Table 3: Gender Distribution

		Frequency	Percent	Valid Percent
Valid	Female	197	55.3	
	Male	159	44.7	
	Total	356	100.0	

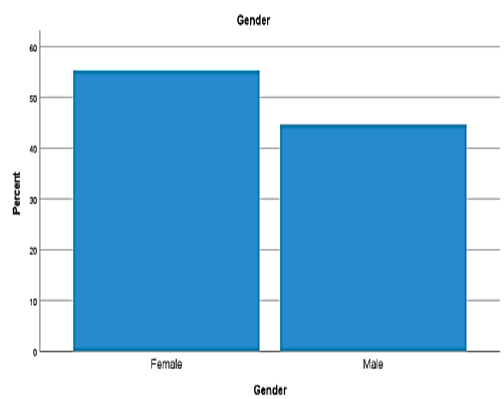

Figure 1: Gender Distribution

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 4, details participants' self-assessed blood sugar control. The largest proportion of respondents (34.3%) reported having moderate control, followed by 29.2% who indicated poor control. Only 19.1% of the sample claimed good control, while 7.9% acknowledged very poor control. A notable 9.6% of participants stated they were unaware of their blood sugar status.

Table 4: Blood Sugar Control Levels

	Frequency	Percent	Valid Percent	Cumulative %
Valid Good control	68	19.1	19.1	19.1
Moderate control	122	34.3	34.3	53.4
No idea	34	9.6	9.6	62.9
Poor control	104	29.2	29.2	92.1
Very Poor control	28	7.9	7.9	100.0
Total	356	100.0	100.0	

The bar chart in Figure 2 visually emphasizes this distribution, indicating a significant proportion of participants have suboptimal or unknown blood glucose control, which could elevate their risk of complications like neuropathy. Cumulatively, over 46% of the sample falls into the poor to very poor control category, while approximately half (53.4%) report either good or moderate control. This suggests that although some level of management exists, a significant percentage of individuals may benefit from improved glycemic education, monitoring, and clinical intervention.

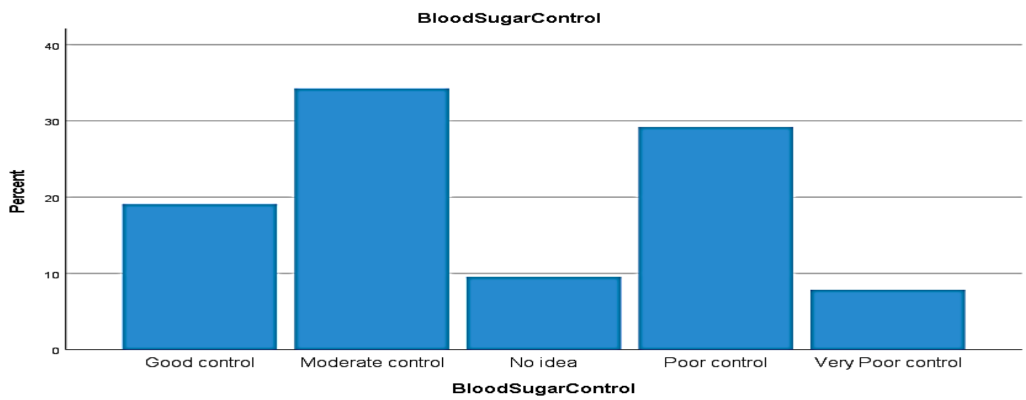

Figure 2: Blood Sugar Control

As shown in Table 5, 206 participants (57.9%) reported having Type 2 diabetes, while 150 (42.1%) had Type 1-diabetes. This reflects the typical population-level distribution, particularly in older adults, where Type 2 diabetes is more common due to age-related insulin resistance, sedentary lifestyles, and dietary patterns.

Table 5: Type of Diabetes

Valid	Frequency	Percent	Valid Percent	Cumulative %
Type 1	150	42.1	42.1	42.1
Type 2	206	57.9	57.9	100.0
Total	356	100.0	100.0	

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Figure 3, visually confirms this trend, reinforcing that Type 2 diabetes dominates the clinical profile of the population under study. Understanding this distinction is important for tailored therapeutic approaches, as the pathophysiology and management strategies differ significantly between the two types.

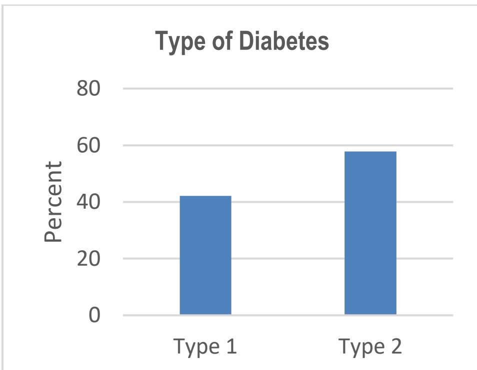

Figure 3: Type of Diabetes

Table 6 indicates that 185 participants (52.0%) reported a positive family history of diabetes, while 171 (48.0%) did not. The near-even split, illustrated in Figure 4.5, underscores the potential genetic and familial component in the development of diabetes among the respondents. A positive family history can serve as a predictor for early screening and preventive strategies, especially among high-risk groups.

Table 6: Family History of Diabetes

	Frequency	Percent	Valid Percent
Valid No	171	48.0	48.0
Yes	185	52.0	52.0
Total	356	100.0	100.0

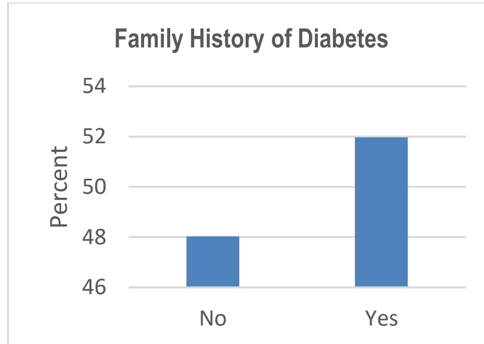

Figure 4: Type of Diabetes

As presented in Table 7, slightly more than half of the participants (50.6%) reported a formal diagnosis of diabetic neuropathy, while 49.4% had not received such a diagnosis. This near-even distribution, visualized in Figure 4, highlights the high prevalence of neuropathy in the study population and underscores the importance of routine screening. Given that diabetic neuropathy often progresses silently, the relatively high diagnosis rate may reflect either heightened clinical awareness or the presence of advanced symptoms in the cohort.

Table 7: Diagnosed DN distribution (Diagnosed Diabetic Neuropathy)

	Frequency	Percent	Valid Percent	Cumulative %
Valid No	176	49.4	49.4	49.4
Yes	180	50.6	50.6	100.0
Total	356	100.0	100.0	

Sensory Symptoms: Tingling, Numbness, and Burning Pain

Tingling sensations, a common early indicator of neuropathy, were reported by 52.0% of respondents, as shown in Table 8 and Figure 5. This symptom's prevalence closely aligns with the diagnosed neuropathy rate, reinforcing its clinical value as an early warning sign.

Table 8: Tingling Sensations

Valid	Frequency	Percent	Valid Percent	Cumulative %
No	171	48.0	48.0	48.0
Yes	185	52.0	52.0	100.0
Total	356	100.0	100.0	

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

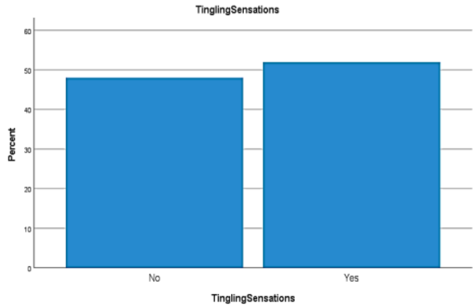

Figure 5: Tingling Sensations Distribution

Numbness, detailed in Table 9 and Figure 6, emerged as the most prevalent symptom, affecting 82.0% of participants. This high proportion suggests widespread sensory nerve involvement and potentially more advanced stages of neuropathy, which increase the risk of complications such as foot ulcers, unnoticed injuries, and infections.

Table 9: Numbness Prevalence

Valid	Frequency	Percent	Valid Percent
No	64	18.0	18.0
Yes	292	82.0	82.0
Total	356	100.0	100.0

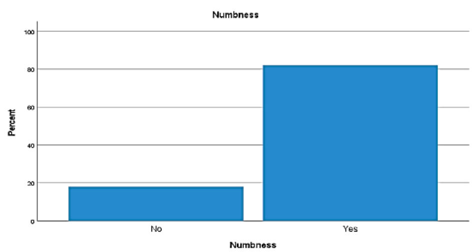

Figure 6: Numbness Distribution

Burning pain, another hallmark symptom of neuropathy, was reported by 53.9% of participants (Table 10, Figure 7). Its relatively high prevalence reflects the chronic discomfort and quality-of-life impairment experienced by individuals with diabetic neuropathy and signals the need for more effective pain management strategies in diabetic care.

Table 10: Burning Pain Prevalence

Valid	Frequency	Percent	Valid Percent
No	164	46.1	46.1
Yes	192	53.9	53.9
Total	356	100.0	100.0

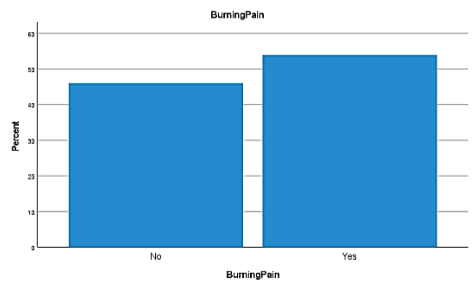

Figure 7: Burning Pain Distribution

A striking 83.7% of respondents (Table 11) reported muscle weakness, indicating that the impact of neuropathy extends beyond sensory deficits to impair motor function. Muscle weakness may stem from peripheral nerve damage, disuse atrophy, or poor glycemic control, and can significantly hinder mobility and independence, especially in older adults.

Table 4.11: Muscle Weakness Prevalence

Valid	Frequency	Percent	Valid Percent	Cumulative %
No	58	16.3	16.3	16.3
Yes	298	83.7	83.7	100.0
Total	356	100.0	100.0	

Hypertension was reported by 52.2% of participants (Table 12, Figure 8); supporting existing evidence linking elevated blood pressure to increased neuropathic risk. Poor vascular health compromises nerve perfusion, accelerating degeneration.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 12: Hypertension Prevalence

Valid	Frequency	Percent	Valid Percent
No	170	47.8	47.8
Yes	186	52.2	52.2
Total	356	100.0	100.0

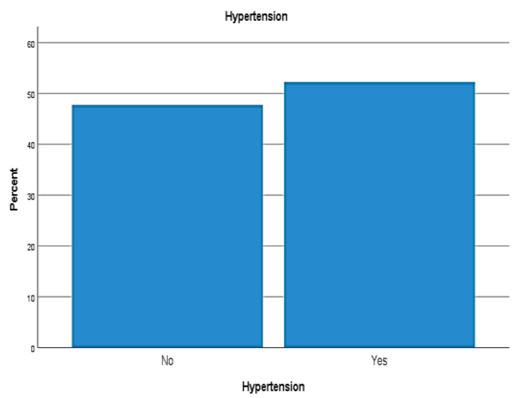

Figure 8: Hypertension Distribution

Dyslipidemia, present in 48.0% of the sample (Table 13, Figure 9), further reinforces the metabolic burden among participants. Dysregulated lipid levels are implicated in oxidative stress and endothelial dysfunction, both of which contribute to neuropathy pathogenesis.

Table 13: Dyslipidemia Prevalence

Valid	Frequency	Percent	Valid Percent
No	185	52.0	52.0
Yes	171	48.0	48.0
Total	356	100.0	100.0

As shown in Table 14, only 4.5% (n=16) of the surveyed individuals reported awareness of the predictive model, while the overwhelming majority (95.5%, n=340) were not aware. This profound lack of awareness underscores a significant communication and educational gap concerning the availability of predictive tools that could enhance early detection and management of diabetic complications. The findings emphasize the urgent need for stakeholder-driven awareness campaigns and the integration of predictive technologies into standard patient care pathways.

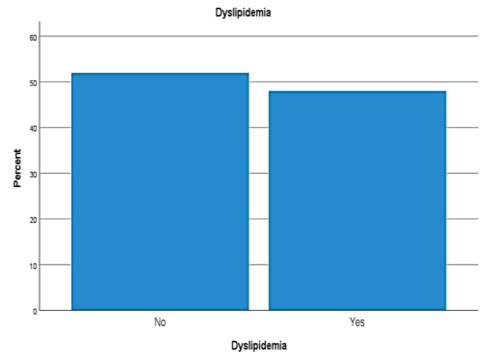

Figure 9: Dyslipidemia Distribution

Table 14: Awareness of Predictive Model

Valid	Frequency	Percent	Valid Percent
No	340	95.5	95.5
Yes	16	4.5	4.5
Total	356	100.0	100.0

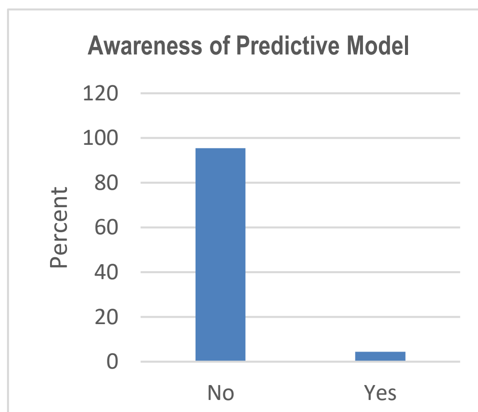

Figure 10: Bar chart summarizing awareness of predictive model.

Table 15, presents the willingness of respondents to adopt machine learning-based predictive tools in managing diabetes. Notably, 90.2% (n=321) of the participants indicated readiness to use such technologies, while only 9.8% (n=35) were reluctant. This readiness is especially remarkable given the low level of awareness reported in Table 4.15. It suggests a high receptivity among patients toward innovative healthcare interventions when properly introduced and explained.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Table 15: Willingness to Use Predictive Model

Valid	Frequency	Percent	Valid Percent
No	35	9.8	9.8
Yes	321	90.2	90.2
Total	356	100.0	100.0

As indicated in Table 16, the majority of participants (82.6%, n=294) believe in the potential of early prediction to mitigate diabetes-related complications, while 17.4% (n=62) held a contrary view. This prevalent belief in proactive care aligns with the high willingness to adopt predictive tools (Table 4.16), reinforcing the acceptability of predictive technologies in clinical settings if awareness barriers are addressed.

Table 16: Belief in Early Prediction

Valid	Frequency	Percent	Valid Percent
No	62	17.4	17.4
Yes	294	82.6	82.6
Total	356	100.0	100.0

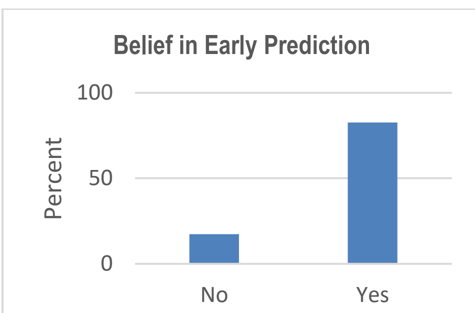

Figure 11: Bar chart summarizing belief in early prediction.

Machine Learning Model Evaluation

Comparative Effectiveness of Machine Learning Models in Predicting Diabetic Neuropathy

This study evaluated the predictive performance of three machine learning algorithms—Naïve Bayes, Support Vector Machine (SVM), and Decision Tree (J48)—in diagnosing diabetic neuropathy using demographic and clinical datasets. The evaluation was based on standard performance metrics: accuracy, precision, recall, and F-measure, each providing a distinct perspective on the models' reliability and diagnostic effectiveness.

The Naïve Bayes classifier yielded an overall accuracy of 83.29%, the lowest among the three models. This outcome reflects the algorithm's simplifying assumption of feature independence, which likely undermined its performance in the presence of complex, interrelated clinical variables. The relatively lower precision and recall scores further suggest that Naïve Bayes struggled to differentiate effectively between diabetic neuropathy-positive and neuropathy-negative cases, thereby reducing its diagnostic robustness.

The Support Vector Machine (SVM**)** demonstrated a markedly higher accuracy of 91.56%, with consistently strong precision and recall scores. This model's performance indicates superior capability in handling nonlinear relationships within the dataset and minimizing classification errors such as false positives and false negatives. The SVM's robustness, especially in high-dimensional and noisy data scenarios, underscores its potential utility for clinical decision support systems targeting diabetic neuropathy.

The Decision Tree (J48) outperformed both Naïve Bayes and SVM, achieving the highest accuracy of 95.06%. Its ability to model complex decision boundaries and account for both categorical and continuous variables made it particularly effective for this dataset. The J48 algorithm also exhibited excellent precision and recall, reflecting its balanced and interpretable classification performance. Moreover, the tree-based structure offers transparency, which is advantageous in clinical applications where explainability is critical.

In summary, while all three models demonstrated varying degrees of effectiveness, the Decision Tree (J48) emerged as the most accurate and reliable algorithm for predicting diabetic neuropathy, followed closely by SVM. The comparatively lower performance of the Naïve Bayes classifier emphasizes the limitations of assuming feature independence in multifactorial medical conditions.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(1), JANUARY, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Key Predictors of Diabetic Neuropathy Identified by the Models

The predictive models identified several key variables that significantly contributed to the accurate classification of diabetic neuropathy. Among the most influential features were clinical symptoms and comorbid conditions commonly associated with neuropathic complications.

Notably, family history of diabetes, persistent pain, numbness, burning sensations, and muscle weakness emerged as prominent indicators, as they are directly linked to nerve damage and sensory impairment. These symptoms served as strong discriminators in distinguishing affected individuals from non-affected counterparts.

Cardiovascular disease, hypertension, obesity, and poor glycemic control were also critical contributors, reflecting the broader metabolic disturbances that exacerbate nerve degeneration. Additionally, advanced age and longer duration of diabetes were shown to elevate neuropathy risk, likely due to cumulative exposure to hyperglycemia and vascular complications.

The presence of retinopathy and chronic kidney disease further signaled progression to advanced diabetic states, where multisystem involvement increases the likelihood of neuropathy. Metabolic indicators such as dyslipidemia and systemic inflammation suggested underlying biochemical imbalances that further predispose patients to nerve injury.

Finally, nerve function test results provided objective, physiological evidence of nerve impairment, making them among the most decisive features across all three models. Their inclusion significantly enhanced model performance by offering quantifiable diagnostic criteria. Together, these variables not only enhanced model accuracy but also provided clinically relevant insights, highlighting the importance of multidimensional assessment in the early identification and management of diabetic neuropathy.

CONCLUSION

This study affirms the potential of machine learning algorithms as effective tools for predicting diabetic neuropathy, thereby facilitating early intervention and mitigating the risk of severe complications. Among the models evaluated, the Decision Tree (J48) demonstrated superior performance, achieving a robust balance between sensitivity and specificity. The identification of significant risk factors—including poor glycemic control, prolonged diabetes duration, hypertension, obesity, and related comorbidities—underscores the necessity of a comprehensive and individualized management strategy. Personalized care plans, tailored to a patient's specific risk profile, are essential for improving clinical outcomes.

Early intervention through appropriate treatment, lifestyle modification, and regular monitoring can substantially reduce the burden of diabetic neuropathy. The findings further highlight the transformative potential of integrating machine learning into routine clinical workflows, offering cost-effective and scalable solutions for chronic disease management. Based on the study's findings, it is recommended that machine learning-based predictive models be integrated into clinical practice to support early diagnosis and personalized care for patients at risk of diabetic neuropathy. Healthcare institutions should also implement training programs to equip professionals with the necessary skills to utilize these tools effectively. Additionally, patient-focused education and awareness campaigns are essential to promote understanding of modifiable risk factors and the importance of early screening. Policymakers are encouraged to support the adoption of data-driven approaches in healthcare planning and chronic disease management to enhance early intervention and improve patient outcomes.

REFERENCES

Abdullah, M., & Abdulazeez, A. M. (2021). Comparative analysis of supervised learning algorithms for predicting diabetes disease. *International Journal of Scientific & Engineering Research, 12*(4), 57–66.

Ahmed, M., Baashar, Y., & Aslam, N. (2022). A hybrid SVM-ANN model for diabetes prediction. *International Journal of*

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

Advanced Computer Science and Applications, 13(6), 112–120.

Akinci, B., Yener, O., Demir, T., & Yesilkaya, E. (2020). Ethnic and gender disparities in diabetic neuropathy prevalence. *Diabetes & Metabolism, 46*(5), 367–374.

Aliyu, U., Ahmed, S. A., & Suleiman, M. (2023). Modifiable risk factors and diabetic complications in northern Nigeria: A cross-sectional study. *Nigerian Journal of Endocrinology, 18*(1), 45–53.

Azeez, M. A., Olatunji, A. A., & Musa, Y. (2022). Epidemiological insights into diabetic neuropathy among African populations. *Pan African Medical Journal, 42*, 223.

Baldimtsi, G., et al. (2024). Age-related progression of diabetic neuropathy: Mechanisms and biomarkers. *Journal of Diabetes and its Complications, 38*, 108459.

Baskozos, G., et al. (2022). Machine learning analysis of pain phenotypes in diabetic peripheral neuropathy. *Pain, 163*(6), e750–e762.

Bhat, A. A., Khan, A., & Farooq, S. (2022). Comparative study of machine learning algorithms for diabetes prediction. *International Journal of Computer Applications, 183*(23), 9–14.

Bhaskar, H., Rajalakshmi, R., & Unnikrishnan, R. (2023). Correlational neural network model for type 2 diabetes prediction using breath biomarkers. *IEEE Journal of Biomedical and Health Informatics, 27*(1), 104–112.

Bondar, R. J., et al. (2021). Pathophysiology and diagnosis of diabetic neuropathy. *Journal of Neurological Disorders, 9*(2), 234–242.

Carmichael, T. R., et al. (2021). Non-invasive biomarkers for diabetic neuropathy detection. *Diabetes Research and Clinical Practice, 179*, 108991.

Chaudhary, S., & Tyagi, R. (2018). Mechanistic understanding of diabetes mellitus. *Journal of Biomedical Sciences, 5*(2), 33–41.

Dubey, S., Meena, Y. K., & Rathi, R. (2020). ANN-based model for diabetic neuropathy risk prediction. *International Journal of Advanced Trends in Computer Science and Engineering, 9*(3), 4101–4106.

El-Massari, A., Salem, H., & Mahfouz, A. (2022). Ontology-enhanced machine learning model for diabetes prediction. *Applied Computing and Informatics, 18*(1), 33–45.

Ergun, U., & Ilhan, A. (2021). Comparison of machine learning classifiers for diabetes prediction. *PeerJ Computer Science, 7*, e673.

Galiero, R., et al. (2023). Shared mechanisms among diabetic microvascular complications. *Frontiers in Endocrinology, 14*, 1054321.

Ibrahim, M. M., & Abdulazeez, A. M. (2021). Hybrid and ensemble machine learning approaches for healthcare prediction. *Journal of Soft Computing and Decision Support Systems, 8*(5), 23–30.

Ismail, N., & Ali, M. (2022). Global trends in diabetes prevalence and mortality. *World Journal of Diabetes, 13*(4), 255–263.

Jain, N., & Devendran, A. (2021). A comprehensive survey on the application of machine learning in healthcare. *International Journal of Engineering Research & Technology, 10*(2), 96–103.

Jijo, A., & Abdulazeez, A. M. (2021). Performance evaluation of decision tree classifiers in medical diagnosis. *Asian Journal of Computer and Information Systems, 9*(4), 67–74.

Khaleel, H., & Al-Bakry, R. (2023). Predictive modeling of diabetes using Pima dataset. *Journal of Computer Applications, 25*(3), 45–52.

Liu, H., Zhang, Q., & Wang, Y. (2022). Oxidative stress markers in diabetic neuropathy. *Oxidative Medicine and Cellular Longevity, 2022*, 3096129.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved

Nascimento, O. J., et al. (2016). Diabetic polyneuropathy: Review of pathophysiology and treatment. *Arquivos de Neuro-Psiquiatria, 74*(5), 366–373.

Olokoba, A. B., et al. (2012). Type 2 diabetes mellitus: A review of current trends. *Oman Medical Journal, 27*(4), 269–273.

Pastore, D., et al. (2022). Dyslipidemia and microvascular damage in diabetic neuropathy. *Metabolism: Clinical and Experimental, 132*, 155203.

Ponirakis, G., et al. (2022). Diabetic neuropathy: Epidemiology, diagnosis, and treatment. *Diabetes/Metabolism Research and Reviews, 38*(6), e3544.

Sadiku, M. N. O., Musa, S. M., & Omotoso, T. (2018). Data mining applications in healthcare. *International Journal of Engineering Research and Advanced Technology, 4*(3), 17–21.

Sanyaolu, A., Okorie, C., & Marinkovic, A. (2023). Global burden of diabetes and health inequality. *Journal of Global Health Reports, 7*, e2023073.

Sisodia, D., & Sisodia, D. S. (2018). Prediction of diabetes using classification algorithms. *Procedia Computer Science, 132*, 1578–1585.

Suryasa, W., Rodriguez-Gamez, M., & Koldoris, T. (2021). Urbanization and lifestyle shifts in global diabetes trends. *Global Public Health Journal, 14*(1), 22–29.

Tiwari, A. K., et al. (2008). Ensemble machine learning approach to tropical diabetic hand syndrome prediction. *Journal of Clinical Bioinformatics, 1*(1), 5–12.

Today Study Group. (2022). Demographic influences on type 2 diabetes outcomes. *New England Journal of Medicine, 387*(4), 322–333.

Wang, Y., Guo, H., & Liu, Q. (2019). Handling imbalanced datasets in diabetes prediction. *Journal of Healthcare Engineering, 2019*, 2356783.

Xu, Y., Wu, J., & Jin, Y. (2020). Predictive modeling for diabetic neuropathy using machine learning. *BMC Medical Informatics and Decision Making, 20*(1), 12.

Zhu, X., Lin, Y., & Luo, Y. (2024). Molecular mechanisms of diabetic peripheral neuropathy. *Frontiers in Molecular Neuroscience, 17*, 1246532.

Corresponding author: Aminu Abdulmajeed Mikailu
abdulmajeed01.am@gmail.com
Alqalam Islamic University, Katsina.
© 2026. Faculty of Technology Education. ATBU Bauchi. All rights reserved