

Original Paper

Use of Automated Machine Learning to Detect Undiagnosed Diabetes in US Adults: Development and Validation Study

Jianxiu Liu^{1,2}, PhD; Fred Ssewamala^{3,4}, PhD; Ruopeng An^{3,5}, MPP, PhD; Mengmeng Ji⁶, MBBS, PhD

¹Division of Sports Science and Physical Education, Tsinghua University, Beijing, China

²IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing, China

³Silver School of Social Work, New York University, New York, NY, United States

⁴McSilver Institute for Poverty Policy and Research, New York University, New York, NY, United States

⁵Constance and Martin Silver Center on Data Science and Social Equity, New York University, New York, NY, United States

⁶Division of Public Health Sciences, Department of Surgery, Washington University in St. Louis, St. Louis, MO, United States

Corresponding Author:

Mengmeng Ji, MBBS, PhD

Division of Public Health Sciences

Department of Surgery, Washington University in St. Louis

600 S Taylor Ave

St. Louis, MO 63110

United States

Phone: 1 2179799336

Email: j.mengmeng@wustl.edu

Abstract

Background: Early diagnosis of diabetes is essential for early interventions to slow the progression of dysglycemia and its comorbidities. However, among individuals with diabetes, about 23% were unaware of their condition.

Objective: This study aims to investigate the potential use of automated machine learning (AutoML) models and self-reported data in detecting undiagnosed diabetes among US adults.

Methods: Individual-level data, including biochemical tests for diabetes, demographic characteristics, family history of diabetes, anthropometric measures, dietary intakes, health behaviors, and chronic conditions, were retrieved from the National Health and Nutrition Examination Survey, 1999–2020. Undiagnosed diabetes was defined as having no prior self-reported diagnosis but meeting diagnostic criteria for elevated hemoglobin A_{1c}, fasting plasma glucose, or 2-hour plasma glucose. The H2O AutoML framework, which allows for automated hyperparameter tuning, model selection, and ensemble learning, was used to automate the machine learning workflow. For comparative analysis, 4 traditional machine learning models—logistic regression, support vector machines, random forest, and extreme gradient boosting—were implemented. Model performance was evaluated using the area under the receiver operating characteristic curve.

Results: The study included 11,815 participants aged 20 years and older, comprising 2256 patients with undiagnosed diabetes and 9559 without diabetes. The average age was 59.76 (SD 15.0) years for participants with undiagnosed diabetes and 46.78 (SD 17.2) years for those without diabetes. The AutoML model demonstrated superior performance compared with the 4 traditional machine learning models. The trained AutoML model achieved an area under the receiver operating characteristic curve of 0.909 (95% CI 0.897–0.921) in the test set. The model demonstrated a sensitivity of 70.26%, specificity of 90.46%, positive predictive value of 64.10%, and negative predictive value of 92.61% for identifying undiagnosed diabetes from nondiabetics.

Conclusions: To our knowledge, this study is the first to utilize the AutoML model for detecting undiagnosed diabetes in US adults. The model's strong performance and applicability to the broader US population make it a promising tool for large-scale diabetes screening efforts.

JMIR AI 2025;4:e68260; doi: [10.2196/68260](https://doi.org/10.2196/68260)

Keywords: machine learning; AutoML; self-report; screening; undiagnosed diabetes

Introduction

Diabetes mellitus is the eighth leading cause of death in the United States and contributes to substantial health care costs [1]. In 2021, an estimated 38.4 million Americans of all ages had diabetes, representing 11.6% of the US population [2,3]. Of those with diabetes, 22.8% were unaware of or did not report having diabetes [2,3]. When diabetes is undiagnosed, and consequently hyperglycemia remains unmanaged, severe and irreversible microvascular and macrovascular complications can develop, including diabetic neuropathy, nephropathy, retinopathy, and cardiovascular disease [4-7].

Screening asymptomatic individuals for undiagnosed diabetes enables earlier diagnosis and treatment, ultimately reducing the risk of complications and premature death [8-10]. The latest American Diabetes Association (ADA) and US Preventive Services Task Force guidelines recommend beginning diabetes screenings at the age of 35 years [11,12]. However, diabetes screening guidelines that rely on blood testing are not widely followed. Only 50%-60% of US adults who met the criteria for screening reported receiving glucose testing within the past 3 years [13,14]. The testing rate was alarmingly low among high-risk groups, including those with low education, low household income, and limited health care access [13,14].

Risk assessment tools for diabetes detection using easily accessible and self-reported data have been proposed, but they have shown low overall accuracy and validity in the general population [15-21]. In recent years, various machine learning algorithms have been used to predict diabetes and have yielded better performance than traditional statistics-based models [22-28]. Few studies have developed machine learning models to detect undiagnosed diabetes in the US population [29,30]. Although 2 studies reported a good overall accuracy of 80%, the quality of a positive prediction by models (ie, precision) was notably low, which could lead to a high number of false positives and unnecessary follow-up testing [29,30].

More recently, there has been growing interest within the health care community in automated machine learning (AutoML), which automates machine learning models' selection, composition, and parameterization to optimize performance [31-33]. AutoML uses voting and stacking ensemble techniques to combine multiple learning models, often improving classification accuracy more effectively than a single machine learning algorithm. Its automation also reduces human error and bias by impartially exploring a wide

range of machine learning models [33]. However, despite its potential, no prior studies have investigated the feasibility and performance of AutoML in screening for undiagnosed diabetes.

This study aimed to investigate the potential use of AutoML and self-reported data in detecting undiagnosed diabetes among US adults in a nationally representative survey. The trained model could aid in detecting undiagnosed diabetes in the general US population, particularly in underserved populations with limited access to blood glucose tests. This study could also promote the adoption of AutoML in diabetes research.

Methods

Data Source

Individual-level data were retrieved from the National Health and Nutrition Examination Survey (NHANES), 1999-2020. NHANES is a nationally representative, repeated cross-sectional study conducted by the National Center for Health Statistics. NHANES adopts a complex, multistage probability sampling design to ensure that the collected data are representative of the noninstitutionalized civilian population in the United States. NHANES includes clinical examinations, selected medical and laboratory tests, and self-reported data. NHANES interviews people in their homes and conducts health examinations in a mobile examination center, including laboratory analysis of blood, urine, and other tissue samples. The detailed study design and methodology of NHANES have been described elsewhere [34,35]. This study followed the CREMLS (Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Models) [36,37].

Biochemical Tests for Undiagnosed Diabetes

Following the ADA guidelines [38], diabetes was diagnosed based on elevated levels of hemoglobin A_{1c} ($\geq 6.5\%$), fasting plasma glucose (≥ 126 mg/dL), or 2-hour plasma glucose (≥ 200 mg/dL) during a 75-g oral glucose tolerance test. In this analysis, undiagnosed diabetes was defined as having no prior self-reported diagnosis but meeting any of the diagnosis criteria for elevated hemoglobin A_{1c}, fasting plasma glucose level, or oral glucose tolerance test level. The details of the diagnostic method used to define diabetes in this study are provided in Table 1.

Table 1. Diagnostic method used to define diabetes in this study.

Diagnostic method	
Diabetic	
Diagnosed diabetes	Answer “Yes” to “Other than during pregnancy, have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?”
Undiagnosed diabetes	Answer “No” to “Other than during pregnancy, have you ever been told by a doctor or health professional that you have diabetes or sugar diabetes?”
AND	

disease, weak/failing kidneys, and cancer/malignancy of any kind.

AutoML and Custom Machine Learning Models

The H2O AutoML framework was used in this study to automate the machine learning workflow. The H2O AutoML trains several models, cross-validated by default, by using the following available algorithms: extreme gradient boosting (XGBoost), gradient boosting machine, generalized linear model, distributed random forest, extremely randomized trees, and fully connected deep neural network. H2O AutoML introduces two essential advancements to optimize model performance. First, it fine-tunes base models using a fast random search approach, where hyperparameters are selected from a range of values identified as most impactful. Second, H2O AutoML leverages a sophisticated stacking technique to create two powerful ensemble models: “All models ensemble,” which combines all the base models trained, and “Best of the Family ensemble,” which contains the best-performing models. The stacked ensemble models are designed to leverage the diverse strengths of various algorithms, resulting in a final model that is accurate and generalizable across different datasets. H2O AutoML has built-in functionality for class balancing and handling of missing values. Detailed documentation, as well as directions for algorithms and the implementation of H2O.ai, are available online [40].

We randomly split the dataset into training (70%) and test (30%) sets. AutoML trained diverse base models on the training set, ranking them by cross-validated area under the receiver operating characteristic curve (AUC). Cross-validation was performed using a 5-fold approach, with models iteratively trained on 4 subsets of the training set and validated on the remaining subset. A stacked ensemble (“leader”) was constructed by blending the top-performing base models via a meta learner. This stacked ensemble model was exported and applied to our independent holdout test set to generate class-probability predictions. The 95% CIs for each AUC were derived from 1000 bootstrap resamples of the test set. Confusion matrices were constructed to calculate sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) using the classification threshold that maximized the F_1 -score that yields the highest harmonic mean of precision and recall on the test set. We extracted feature importance from each base learner and computed a weighted aggregate in the H2O leader model. For tree-based models, feature importance was determined by the frequency of each feature used for splitting and the overall reduction in squared error. For non-tree-based models, importance was based on coefficient magnitudes.

In addition to the AutoML model, we conducted a comparative analysis using 4 traditional machine learning models—logistic regression, support vector machines, random forest, and XGBoost, with Synthetic Minority Over-sampling Technique applied to address data imbalance during training [41,42].

The clinical guidelines generally recommend confirming an elevated test with a secondary measurement for the diagnosis of diabetes [38,43]. In the main analyses, diabetes was diagnosed based on a single elevated test result. In additional analyses, the diagnosis of diabetes was confirmed by at least 2 elevated tests recommended by the ADA guidelines. We also evaluated the performance of a 3-class prediction model for diabetes within the AutoML framework, using classification schemes that distinguished between normoglycemia, prediabetes, and undiagnosed diabetes.

Summary statistics for participant characteristics, stratified by diabetes status, were calculated. Categorical variables were compared using χ^2 tests, and continuous variables were evaluated using independent samples t tests. Missing data were imputed using mean values for continuous variables and mode for categorical variables. We used STATA 18 (StataCorp LLC) for data preparation and Python version 3.10.12 (Python Software Foundation) to implement the H2O AutoML (H2O.ai, Inc) and custom machine learning models (version 3.46.0.6).

Ethical Considerations

This study used publicly available, deidentified NHANES data. In accordance with the US Department of Health and Human Services (Title 45 of the Code of Federal Regulations; §46.104 (d), section 4) [44], analyses of publicly available, deidentified data are not considered human subjects research and therefore do not require review by the Washington University Institutional Review Board.

Results

In total, the study cohort included 11,815 participants, with 9559 participants without diabetes and 2256 with undiagnosed diabetes. The characteristics of the study cohort are summarized in Table 2. The average ages were 59.76 (SD 15.0) years for those with undiagnosed diabetes and 46.78 (SD 17.2) years for those without. The diabetes group had a higher proportion of males, lower levels of education, and a greater likelihood of having a relative with diabetes. Additionally, patients with undiagnosed diabetes had higher BMI, larger waist circumference, and a higher prevalence of chronic conditions. The flow diagram of participant selection is presented in Figure 1.

Table 2. Cohort characteristics by diabetes status.

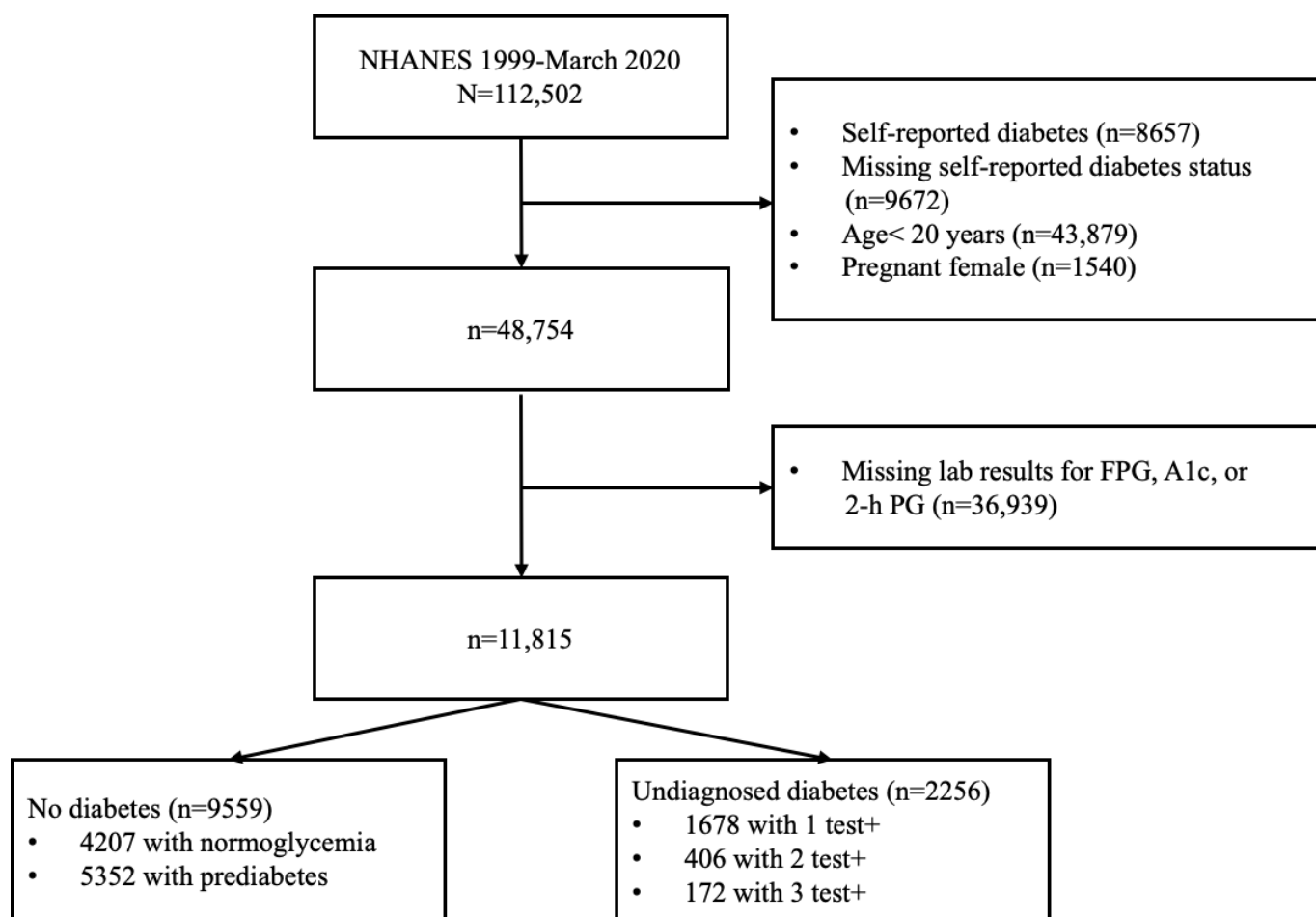
	No diabetes (n=9559)	Undiagnosed diabetes (n=2256)	P value
Laboratory tests, mean (SD)			
Glycohemoglobin (%)	5.41 (0.38)	6.88 (1.60)	<.001

	No diabetes (n=9559)	Undiagnosed diabetes (n=2256)	P value
Fasting glucose plasma (mg/dL)	98.18 (9.46)	141.38 (46.79)	<.001
Oral glucose tolerance test (mg/dL)	109.53 (31.95)	229.24 (75.68)	<.001
Demographic characteristics			
Age (years), mean (SD)	46.78 (17.19)	59.76 (14.98)	<.001
Gender, n (%)			.005
Male	4711 (49.28)	1186 (52.57)	
Female	4848 (50.72)	1070 (47.43)	
Race/ethnicity, n (%)			<.001
Non-Hispanic White	4381 (45.83)	835 (37.01)	
Non-Hispanic Black	1750 (18.31)	504 (22.34)	
Hispanic	2459 (25.72)	698 (30.94)	
Other	969 (10.14)	219 (9.71)	
Education, n (%)			<.001
≤9 grade	881 (9.22)	434 (19.28)	
9th-11th grade	1301 (13.62)	391 (17.37)	
High school	2127 (22.27)	554 (24.61)	
Some college	2801 (29.32)	543 (24.12)	
College or above	2443 (25.57)	329 (14.62)	
Marital status, n (%)			<.001
Married	5016 (52.47)	1248 (55.54)	
Widowed	568 (5.94)	390 (17.36)	
Divorced	950 (9.94)	251 (11.17)	
Separated	330 (3.45)	72 (3.20)	
Never married	1843 (19.28)	192 (8.54)	
Living with partner	852 (8.91)	94 (4.18)	
Income-to-poverty ratio, mean (SD)	2.60 (1.64)	2.29 (1.53)	<.001
Family history of diabetes, n (%)	3394 (36.16)	1072 (48.79)	<.001
Anthropometric measures, mean (SD)			
Height (cm)	167.89 (10.00)	165.84 (10.24)	<.001
Weight (kg)	80.12 (20.69)	88.19 (23.32)	<.001
BMI (kg/m ²)	28.33 (6.53)	31.94 (7.43)	<.001
Waist circumference (cm)	96.92 (15.32)	107.5 (15.69)	<.001
Diet intake and eating behavior, mean (SD)			
Daily total energy (kcal)	2090.13 (837.69)	1912.17 (851.19)	<.001
Daily total fat (g)	78.27 (38.36)	72.29 (39.76)	<.001
Daily total sugars (g)	113.59 (65.37)	104.57 (68.14)	<.001
Daily total cholesterol (mg)	287.05 (193.07)	285.82 (201.17)	.80
Times of dining out per week	3.43 (3.84)	2.60 (3.43)	<.001
Health behaviors			
Physical activity (MET ^a -h/wk), mean (SD)	2.36 (3.69)	1.62 (2.81)	<.001
Smoking, n (%)	4200 (43.97)	1031 (48.91)	<.001
Drinking, n (%)	6758 (74.26)	1262 (67.52)	<.001
Self-reported chronic conditions, n (%)			
Hypertension	2778 (29.10)	1188 (52.85)	<.001
Rheumatoid arthritis	2206 (23.12)	778 (34.55)	<.001
Myocardial infarction	283 (2.96)	152 (6.76)	<.001
Congestive heart failure	182 (1.91)	112 (4.99)	<.001

	No diabetes (n=9559)	Undiagnosed diabetes (n=2256)	<i>P</i> value
Coronary heart disease	261 (2.74)	145 (6.47)	<.001
Stroke	231 (2.42)	127 (5.64)	<.001
Liver disease	295 (3.09)	117 (5.19)	<.001
Weak/failing kidneys	199 (2.08)	64 (2.84)	.03
Cancer	760 (7.95)	257 (11.40)	<.001

^aMET: metabolic equivalent.

Figure 1. Flow diagram of participant selection. 2-h PG: 2-hour plasma glucose; A_{1c}: hemoglobin A_{1c}; FPG: fasting plasma glucose; NHANES: National Health and Nutrition Examination Survey.



The performance of the AutoML model and traditional machine learning models is summarized in [Figure 2](#). The AutoML model demonstrated superior performance compared to the 4 traditional machine learning models—logistic regression, support vector machines, random forest, and XGBoost. The trained AutoML model achieved an AUC

of 0.909 (95% CI 0.897-0.921) and an accuracy of 86.5% in the test set. The model demonstrated a sensitivity of 70.26%, specificity of 90.46%, PPV of 64.10%, and NPV of 92.61% for identifying undiagnosed diabetes from nondiabetics ([Table 3](#)). The model summary and details were provided in [Multimedia Appendix 1](#).

Figure 2. Performance of AutoML model and custom machine learning models in detecting undiagnosed diabetes on the test set. AutoML: automated machine learning; XGBoost: extreme gradient boosting.

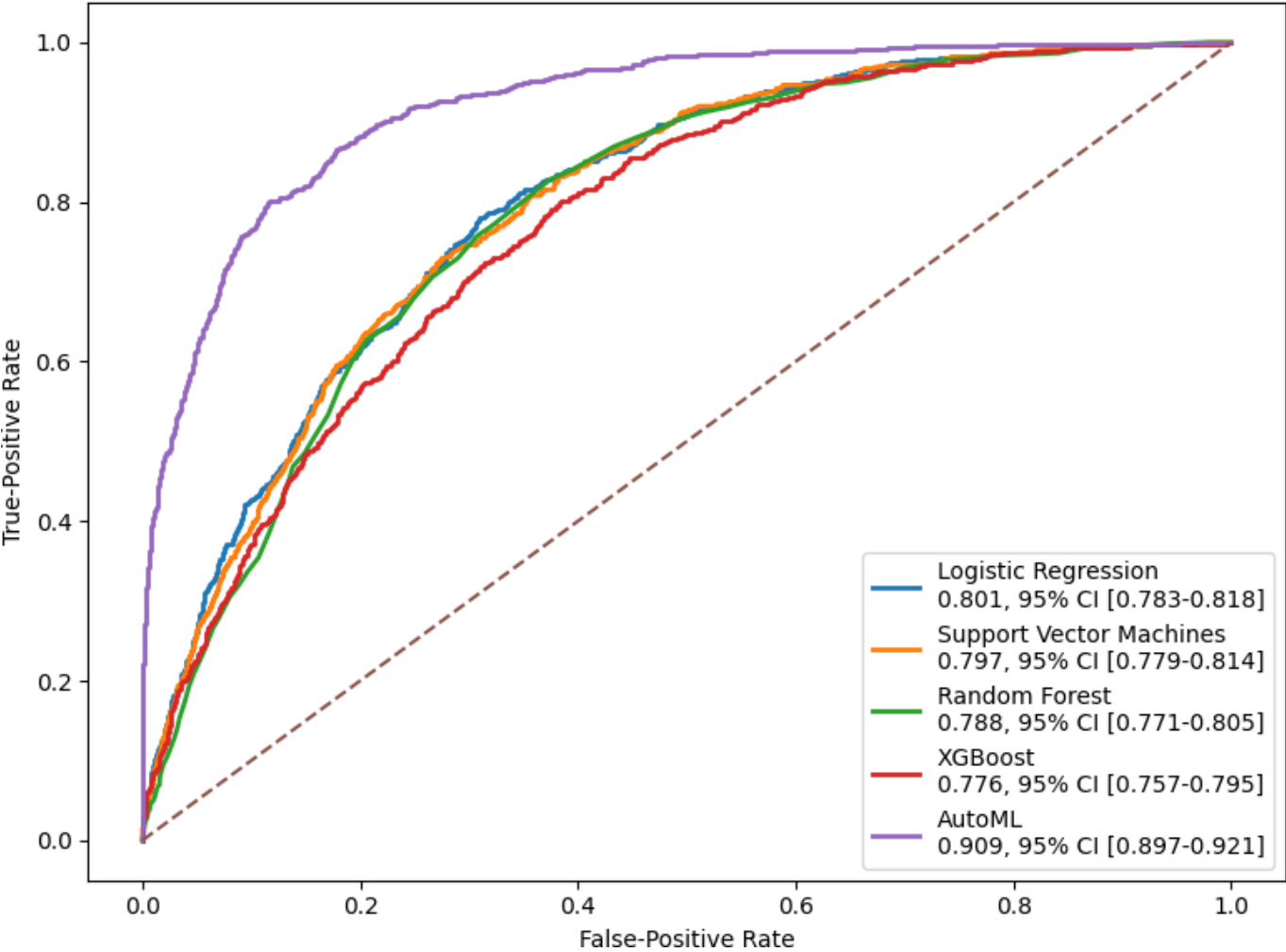


Table 3. Confusion matrix for the classification of undiagnosed diabetes using the AutoML model.^a

True label	Predicted label, n (%)	
	No diabetes	Undiagnosed diabetes
No diabetes	2558 (73)	270 (8)
Undiagnosed diabetes	204 (6)	482 (14)

^aNote: The cutoff threshold was 0.248718, optimized for F_1 -score that maximized the harmonic mean of precision and recall. The matrix shows the distribution of true labels against predicted labels. The cell values indicate the number of instances (absolute counts) and their corresponding percentages of the total in the test data. Sensitivity (482/686, 70.26%), specificity (2558/2828, 90.45%), positive predictive value (482/752, 64.10%), and negative predictive value (2558/2762, 92.61%) are derived from the matrix to assess model performance.

The top 5 features are age, waist circumference, daily total sugar intake, income, and BMI, together accounting for 50% of total model importance. Comorbidities, except hypertension, contributed minimally relative to demographic and behavioral factors. Excluding comorbidities (except hypertension) resulted in comparable model performance, with an AUC of 0.830 and an accuracy of 85.1%.

Additional analysis results are summarized in Table 4. The model using diabetes diagnosis criteria with a confirmatory test achieved a testing accuracy of 98.0% and 89.7% with

an AUC of 0.823. However, precision (ie, PPV) and recall (ie, sensitivity) were suboptimal due to the small number of patients meeting the diabetes criteria with ≥ 2 tests. The model demonstrated a sensitivity of 44.10%, specificity of 92.22%, PPV of 24.23%, and NPV of 96.69% for identifying undiagnosed diabetes from nondiabetes. The performance of the multiclass prediction model was poor, with an overall accuracy of 58.9% for using diagnosis criteria with ≥ 1 test and 67.1% for using diagnosis criteria with ≥ 2 tests.

Table 4. Additional models for alternative diabetes diagnosis criteria with the second confirmative test and multiclass prediction including prediabetes.

	Undiagnosed diabetes (≥ 2 test) versus no diabetes ^a		Undiagnosed diabetes (≥ 1 test) versus prediabetes versus normoglycemia		Undiagnosed diabetes (≥ 2 test) versus prediabetes versus normoglycemia	
	Train	Test	Train	Test	Train	Test
Accuracy (%)	98.0	89.7	66.8	59.0	68.6	67.1
Overall AUC ^b	0.993	0.823	0.627	0.557	0.588	0.557
AUC (normal versus rest)	— ^c	—	0.502	0.445	0.316	0.307
AUC (prediabetes versus rest)	—	—	0.814	0.652	0.792	0.715
AUC (diabetes versus rest)	—	—	0.415	0.545	0.555	0.673

^a“No diabetes” group includes normoglycemia and prediabetes.

^bAUC: area under the receiver operating characteristic curve.

^cNot applicable.

Discussion

To our knowledge, this study is the first to utilize the AutoML model for detecting undiagnosed diabetes in US adults. The best-performing model achieved an AUC of 0.91 and an accuracy of 86.5% in the test set. National surveillance shows that nearly half of those with undiagnosed diabetes have hypertension, lipid abnormalities, or cardiovascular and chronic kidney diseases [45–48]. Delayed diagnosis of diabetes hinders the opportunities for early intervention to slow the progression of dysglycemia and its comorbidities. Our model was trained and tested using a substantial and diverse dataset comprising nationally representative survey data. The model’s high accuracy and applicability to the broader US population make it a promising tool for large-scale diabetes screening efforts.

The feature importance ranking of the best-performing model highlights waist circumference, BMI, and dietary variables as key predictors, underscoring their strong links to metabolic health, insulin resistance, and dietary habits that influence diabetes risk. Lifestyle factors such as drinking frequency and physical activity also emerged as significant contributors, whereas self-reported comorbidities played a smaller role once anthropometric and behavioral measures were included. These findings align with epidemiological studies and improve the model’s interpretability, providing actionable insights to prioritize targeted interventions for modifiable risk factors [49].

Machine learning has advanced clinical research but faces adoption barriers like data access, imbalances, and reliance on data science expertise for deployment [50,51]. AutoML reduces the need for machine learning expertise, enabling clinicians to use advanced technologies without programming skills and integrating them into research and clinical practice [52–55]. Despite being promising, few studies have explored the application of AutoML for diabetes diagnosis [56].

Previous studies have compared traditional machine learning models with conventional statistical models for identifying undiagnosed diabetes, demonstrating that machine

learning models outperform statistical models [29,30]. These studies reported AUC values between 0.73 and 0.81, consistent with the performance of traditional models in this study. However, the reported low PPVs highlighted the limited ability of these models to accurately identify undiagnosed diabetes cases [30]. This study showed that AutoML models are superior and outperformed traditional machine learning models in detecting undiagnosed diabetes. Similarly, one study has reported that the AutoML model outperformed both individual and ensemble models in identifying patients with diabetes using electronic medical records data [57]. These findings suggest that AutoML provides a more accessible and efficient approach, eliminating the need for manual optimization while delivering superior performance.

Nonetheless, several issues with the AutoML model in diabetes screening should be noted. When applying more stringent diabetes diagnostic criteria, the accuracy reached 90%; however, precision and recall were low, likely due to the limited number of samples that met the ≥ 2 test criteria. H2O AutoML provides significant advantages, including its built-in class balancing functionality, which automates the handling of moderate class imbalance without requiring external implementations. Using random sampling to upsample minority classes or downsample majority classes, it effectively manages datasets with moderate imbalance, such as the 4:1 ratio under the 1+ test criterion for diabetes diagnosis in this study. However, with the stricter 2+ test criterion, the imbalance ratio rose to 24:1, likely exceeding H2O AutoML’s ability to mitigate the imbalance. This severe imbalance impacted the model’s ability to accurately distinguish undiagnosed diabetes from nondiabetes in unseen data, reducing precision and recall. For such highly imbalanced datasets, combining AutoML with other data resampling methods, such as Synthetic Minority Over-sampling Technique, could better improve model performance [58].

In addition, the model’s performance in multiclass classification of no diabetes, prediabetes, and undiagnosed diabetes was notably poor. Prediabetes is an intermediate stage between normal glycemia and diabetes and is highly

prevalent [59]. Clinically, prediabetes and diabetes, as the continuum of dysglycemia, share many overlapping risk factors, such as insulin resistance and elevated glucose levels [60], making their differentiation challenging. The subtle metabolic differences between these conditions may not have been adequately captured by the included self-reported data. Future efforts should focus on incorporating additional features and refining the model architecture to enhance accuracy and improve its ability to identify prediabetes.

A major strength of this model is its use of NHANES data, which is nationally representative of the US population, enhancing the generalizability of the findings. The model incorporates comprehensive self-reported data, including nutritional information, for predicting undiagnosed diabetes. The application of AutoML in this study represents the first use of this approach in diabetes research, providing a foundation for further development and validation of similar models. However, the widespread adoption of our AutoML in health care requires further development and validation. False positives can lead to unnecessary tests, higher costs, and patient anxiety, while false negatives may delay treatment and worsen outcomes. Using AutoML models in real-world clinical settings requires meticulous threshold optimization and validation to achieve an appropriate balance between precision and recall. These models should be rigorously evaluated for fairness and equity, ensuring that performance does not vary significantly across demographic groups [61, 62]. Practical barriers to implementing AutoML in clinical practice also include integration with electronic health record

systems and the lack of trust in “black-box” models due to their opacity [31,61-63].

Several limitations should be noted. First, the definition of undiagnosed diabetes in the base model was based on a single elevated measurement, which may not fully capture the condition. However, the model still showed utility as a screening tool. Second, the model demonstrated poor performance in multiclass prediction, including prediabetes, indicating that additional feature refinement and model adjustments may be necessary to improve accuracy in 3-class predictions. Third, known diabetes diagnoses rely on self-reports, which may introduce potential recall bias and misclassification. This could lead to mislabeling cases and diminishing model performance. Although most published public health studies still rely on self-reports—hence making our study not an unusual one in using self-reports—future studies should aim to validate self-reported data against medical records where feasible to minimize errors and increase reliability of the observations. Finally, the generalizability of this model may be limited to US populations and may not extend to non-US populations.

This study demonstrates the potential of AutoML in detecting undiagnosed diabetes using self-reported and easily accessible data. Although challenges remain in accurately classifying multiple categories, including prediabetes, the model shows promise as a tool for large-scale diabetes screening. Further refinement and validation are required to improve its applicability across diverse populations.

Acknowledgments

JL was supported by Beijing Educational Sciences planning project (CCHA241) and Sports Special Project of China Disabled Persons' Federation (2024 KFJS & 001).

Authors Contributions

MJ and RA conceived and designed the study. JL and MJ contributed to data curation, formal analysis, funding acquisition, and investigation. MJ was responsible for resources and validation. Visualization was performed by JL and MJ. The original draft of the manuscript was written by JL and MJ, with critical review and editing provided by JL, FS, RA, and MJ.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Model summary for stacked ensemble.

[DOCX File (Microsoft Word File), 16 KB-Multimedia Appendix 1]

References

1. Ahmad FB, Anderson RN. The leading causes of death in the US for 2020. JAMA. May 11, 2021;325(18):1829-1830. [doi: [10.1001/jama.2021.5469](https://doi.org/10.1001/jama.2021.5469)] [Medline: [33787821](https://pubmed.ncbi.nlm.nih.gov/33787821/)]
2. Statistics about diabetes. American Diabetes Association. URL: <https://diabetes.org/about-diabetes/statistics/about-diabetes> [Accessed 2025-01-27]
3. Diabetes statistics. National Institute of Diabetes and Digestive and Kidney Diseases. URL: <https://www.niddk.nih.gov/health-information/health-statistics/diabetes-statistics> [Accessed 2025-01-27]
4. Beagley J, Guariguata L, Weil C, Motala AA. Global estimates of undiagnosed diabetes in adults. Diabetes Res Clin Pract. Feb 2014;103(2):150-160. [doi: [10.1016/j.diabres.2013.11.001](https://doi.org/10.1016/j.diabres.2013.11.001)] [Medline: [24300018](https://pubmed.ncbi.nlm.nih.gov/24300018/)]
5. Fowler MJ. Microvascular and macrovascular complications of diabetes. Clin Diabetes. Jul 1, 2011;29(3):116-122. [doi: [10.2337/diaclin.29.3.116](https://doi.org/10.2337/diaclin.29.3.116)]

6. Deshpande AD, Harris-Hayes M, Schootman M. Epidemiology of diabetes and diabetes-related complications. *Phys Ther.* Nov 2008;88(11):1254-1264. [doi: [10.2522/ptj.20080020](https://doi.org/10.2522/ptj.20080020)] [Medline: [18801858](https://pubmed.ncbi.nlm.nih.gov/18801858/)]
7. Ahmad F, Joshi SH. Self-care practices and their role in the control of diabetes: a narrative review. *Cureus.* Jul 2023;15(7):e41409. [doi: [10.7759/cureus.41409](https://doi.org/10.7759/cureus.41409)] [Medline: [37546053](https://pubmed.ncbi.nlm.nih.gov/37546053/)]
8. Ali MK, McKeever Bullard K, Imperatore G, et al. Reach and use of diabetes prevention services in the United States, 2016-2017. *JAMA Netw Open.* May 3, 2019;2(5):e193160. [doi: [10.1001/jamanetworkopen.2019.3160](https://doi.org/10.1001/jamanetworkopen.2019.3160)] [Medline: [31074808](https://pubmed.ncbi.nlm.nih.gov/31074808/)]
9. Ali MK, Bullard KM, Gregg EW, Del Rio C. A cascade of care for diabetes in the United States: visualizing the gaps. *Ann Intern Med.* Nov 18, 2014;161(10):681-689. [doi: [10.7326/M14-0019](https://doi.org/10.7326/M14-0019)] [Medline: [25402511](https://pubmed.ncbi.nlm.nih.gov/25402511/)]
10. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet.* Nov 14, 2009;374(9702):1677-1686. [doi: [10.1016/S0140-6736\(09\)61457-4](https://doi.org/10.1016/S0140-6736(09)61457-4)] [Medline: [19878986](https://pubmed.ncbi.nlm.nih.gov/19878986/)]
11. US Preventive Services Task Force, Davidson KW, Barry MJ, et al. Screening for prediabetes and type 2 diabetes: US Preventive Services Task Force recommendation statement. *JAMA.* Aug 24, 2021;326(8):736-743. [doi: [10.1001/jama.2021.12531](https://doi.org/10.1001/jama.2021.12531)] [Medline: [34427594](https://pubmed.ncbi.nlm.nih.gov/34427594/)]
12. American Diabetes Association Professional Practice Committee. Classification and diagnosis of diabetes: standards of medical care in diabetes—2022. *Diabetes Care.* Jan 1, 2022;45(Supplement_1):S17-S38. [doi: [10.2337/dc22-S002](https://doi.org/10.2337/dc22-S002)]
13. Bullard KM, Ali MK, Imperatore G, et al. Receipt of glucose testing and performance of two US diabetes screening guidelines, 2007-2012. *PLoS ONE.* 2015;10(4):e0125249. [doi: [10.1371/journal.pone.0125249](https://doi.org/10.1371/journal.pone.0125249)] [Medline: [25928306](https://pubmed.ncbi.nlm.nih.gov/25928306/)]
14. Ali MK, Imperatore G, Benoit SR, et al. Impact of changes in diabetes screening guidelines on testing eligibility and potential yield among adults without diagnosed diabetes in the United States. *Diabetes Res Clin Pract.* Mar 2023;197:110572. [doi: [10.1016/j.diabres.2023.110572](https://doi.org/10.1016/j.diabres.2023.110572)] [Medline: [36775024](https://pubmed.ncbi.nlm.nih.gov/36775024/)]
15. Franciosi M, De Berardis G, Rossi MCE, et al. Use of the diabetes risk score for opportunistic screening of undiagnosed diabetes and impaired glucose tolerance: the IGLOO (Impaired Glucose Tolerance and Long-Term Outcomes Observational) study. *Diabetes Care.* May 2005;28(5):1187-1194. [doi: [10.2337/diacare.28.5.1187](https://doi.org/10.2337/diacare.28.5.1187)] [Medline: [15855587](https://pubmed.ncbi.nlm.nih.gov/15855587/)]
16. Lee YH, Bang H, Kim HC, Kim HM, Park SW, Kim DJ. A simple screening score for diabetes for the Korean population: development, validation, and comparison with other scores. *Diabetes Care.* Aug 2012;35(8):1723-1730. [doi: [10.2337/dc11-2347](https://doi.org/10.2337/dc11-2347)] [Medline: [22688547](https://pubmed.ncbi.nlm.nih.gov/22688547/)]
17. Bang H, Edwards AM, Bomback AS, et al. Development and validation of a patient self-assessment score for diabetes risk. *Ann Intern Med.* Dec 1, 2009;151(11):775-783. [doi: [10.7326/0003-4819-151-11-200912010-00005](https://doi.org/10.7326/0003-4819-151-11-200912010-00005)] [Medline: [19949143](https://pubmed.ncbi.nlm.nih.gov/19949143/)]
18. Herman WH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE. A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. *Diabetes Care.* Mar 1995;18(3):382-387. [doi: [10.2337/diacare.18.3.382](https://doi.org/10.2337/diacare.18.3.382)] [Medline: [7555482](https://pubmed.ncbi.nlm.nih.gov/7555482/)]
19. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes risk calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes Care.* May 2008;31(5):1040-1045. [doi: [10.2337/dc07-1150](https://doi.org/10.2337/dc07-1150)] [Medline: [18070993](https://pubmed.ncbi.nlm.nih.gov/18070993/)]
20. Rolka DB, Narayan KM, Thompson TJ, et al. Performance of recommended screening tests for undiagnosed diabetes and dysglycemia. *Diabetes Care.* Nov 2001;24(11):1899-1903. [doi: [10.2337/diacare.24.11.1899](https://doi.org/10.2337/diacare.24.11.1899)] [Medline: [11679454](https://pubmed.ncbi.nlm.nih.gov/11679454/)]
21. Dhippayom T, Chaiyakunapruk N, Krass I. How diabetes risk assessment tools are implemented in practice: a systematic review. *Diabetes Res Clin Pract.* Jun 2014;104(3):329-342. [doi: [10.1016/j.diabres.2014.01.008](https://doi.org/10.1016/j.diabres.2014.01.008)] [Medline: [24485859](https://pubmed.ncbi.nlm.nih.gov/24485859/)]
22. Maniruzzaman M, Rahman MJ, Al-Mehedi Hasan M, et al. Accurate diabetes risk stratification using machine learning: role of missing value and outliers. *J Med Syst.* Apr 10, 2018;42(5):92. [doi: [10.1007/s10916-018-0940-7](https://doi.org/10.1007/s10916-018-0940-7)] [Medline: [29637403](https://pubmed.ncbi.nlm.nih.gov/29637403/)]
23. Hasan MK, Alam MA, Das D, Hossain E, Hasan M. Diabetes prediction using ensembling of different machine learning classifiers. *IEEE Access.* 2020;8:76516-76531. [doi: [10.1109/ACCESS.2020.2989857](https://doi.org/10.1109/ACCESS.2020.2989857)] [Medline: [34812373](https://pubmed.ncbi.nlm.nih.gov/34812373/)]
24. Sisodia D, Sisodia DS. Prediction of diabetes using classification algorithms. *Procedia Comput Sci.* 2018;132:1578-1585. [doi: [10.1016/j.procs.2018.05.122](https://doi.org/10.1016/j.procs.2018.05.122)]
25. Olivera AR, Roesler V, Iochpe C, et al. Comparison of machine-learning algorithms to build a predictive model for detecting undiagnosed diabetes - ELSA-Brasil: accuracy study. *Sao Paulo Med J.* 2017;135(3):234-246. [doi: [10.1590/1516-3180.2016.0309010217](https://doi.org/10.1590/1516-3180.2016.0309010217)] [Medline: [28746659](https://pubmed.ncbi.nlm.nih.gov/28746659/)]
26. Ryu KS, Lee SW, Batbaatar E, Lee JW, Choi KS, Cha HS. A deep learning model for estimation of patients with undiagnosed diabetes. *Appl Sci (Basel).* 2020;10(1):421. [doi: [10.3390/app10010421](https://doi.org/10.3390/app10010421)]
27. Choi SG, Oh M, Park DH, et al. Comparisons of the prediction models for undiagnosed diabetes between machine learning versus traditional statistical methods. *Sci Rep.* Aug 11, 2023;13(1):13101. [doi: [10.1038/s41598-023-40170-0](https://doi.org/10.1038/s41598-023-40170-0)] [Medline: [37567907](https://pubmed.ncbi.nlm.nih.gov/37567907/)]

28. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. 2017;15:104-116. [doi: [10.1016/j.csbj.2016.12.005](https://doi.org/10.1016/j.csbj.2016.12.005)] [Medline: [28138367](https://pubmed.ncbi.nlm.nih.gov/28138367/)]
29. Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med Inform Decis Mak*. Mar 22, 2010;10(1):16. [doi: [10.1186/1472-6947-10-16](https://doi.org/10.1186/1472-6947-10-16)] [Medline: [20307319](https://pubmed.ncbi.nlm.nih.gov/20307319/)]
30. Cichosz SL, Bender C, Hejlesen O. A comparative analysis of machine learning models for the detection of undiagnosed diabetes patients. *Diabetology*. 2024;5(1):1-11. [doi: [10.3390/diabetology5010001](https://doi.org/10.3390/diabetology5010001)]
31. Waring J, Lindvall C, Umeton R. Automated machine learning: review of the state-of-the-art and opportunities for healthcare. *Artif Intell Med*. Apr 2020;104:101822. [doi: [10.1016/j.artmed.2020.101822](https://doi.org/10.1016/j.artmed.2020.101822)] [Medline: [32499001](https://pubmed.ncbi.nlm.nih.gov/32499001/)]
32. Scott IA, De Guzman KR, Falconer N, et al. Evaluating automated machine learning platforms for use in healthcare. *JAMIA Open*. Jul 2024;7(2):ooae031. [doi: [10.1093/jamiaopen/ooae031](https://doi.org/10.1093/jamiaopen/ooae031)] [Medline: [38863963](https://pubmed.ncbi.nlm.nih.gov/38863963/)]
33. Luo G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. *Netw Model Anal Health Inform Bioinforma*. Dec 2016;5(1):18. [doi: [10.1007/s13721-016-0125-6](https://doi.org/10.1007/s13721-016-0125-6)]
34. Chen TC, Clark J, Riddles MK, Mohadjer LK, Fakhouri THI. National Health and Nutrition Examination Survey, 2015-2018: sample design and estimation procedures. *Vital Health Stat 2*. Apr 2020(184):1-35. [Medline: [33663649](https://pubmed.ncbi.nlm.nih.gov/33663649/)]
35. Mirel LB, Mohadjer LK, Dohrmann SM, et al. National Health and Nutrition Examination Survey: estimation procedures, 2007-2010. *Vital Health Stat 2*. Aug 2013;2(159):1-17. [Medline: [25093338](https://pubmed.ncbi.nlm.nih.gov/25093338/)]
36. El Emam K, Leung TI, Malin B, Klement W, Eysenbach G. Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Models (CREMLS). *J Med Internet Res*. May 2, 2024;26:e52508. [doi: [10.2196/52508](https://doi.org/10.2196/52508)] [Medline: [38696776](https://pubmed.ncbi.nlm.nih.gov/38696776/)]
37. Klement W, El Emam K. Consolidated reporting guidelines for prognostic and diagnostic machine learning modeling studies: development and validation. *J Med Internet Res*. Aug 31, 2023;25:e48763. [doi: [10.2196/48763](https://doi.org/10.2196/48763)] [Medline: [37651179](https://pubmed.ncbi.nlm.nih.gov/37651179/)]
38. ElSayed NA, Aleppo G, Bannuru RR, et al. Diagnosis and classification of diabetes: standards of care in diabetes—2024. *Diabetes Care*. Jan 1, 2024;47(Supplement_1):S20-S42. [doi: [10.2337/dc24-S002](https://doi.org/10.2337/dc24-S002)]
39. NHANES 2005-2006 data documentation, codebook, and frequencies. Centers for Disease Control and Prevention. URL: https://wwwn.cdc.gov/Nchs/Nhanes/2005-2006/PAQIAF_D.htm [Accessed 2025-01-27]
40. LeDell E, Poirier S. H2O AutoML: scalable automatic machine learning. Presented at: 7th ICML Workshop on Automated Machine Learning. Jul 18, 2020. URL: https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf [Accessed 2025-01-27]
41. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-Sampling Technique. *J Artif Intell Res*. 2002;16:321-357. [doi: [10.1613/jair.953](https://doi.org/10.1613/jair.953)]
42. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*. Mar 22, 2013;14:106. [doi: [10.1186/1471-2105-14-106](https://doi.org/10.1186/1471-2105-14-106)] [Medline: [23522326](https://pubmed.ncbi.nlm.nih.gov/23522326/)]
43. Samson SL, Vellanki P, Blonde L, et al. American Association of Clinical Endocrinology consensus statement: comprehensive type 2 diabetes management algorithm 2023 update. *Endocr Pract*. May 2023;29(5):305-340. [doi: [10.1016/j.eprac.2023.02.001](https://doi.org/10.1016/j.eprac.2023.02.001)] [Medline: [37150579](https://pubmed.ncbi.nlm.nih.gov/37150579/)]
44. Code of federal regulations. US Department of Health and Human Services. URL: <https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.104> [Accessed 2024-9-30]
45. Plantinga LC, Crews DC, Coresh J, et al. Prevalence of chronic kidney disease in US adults with undiagnosed diabetes or prediabetes. *Clin J Am Soc Nephrol*. Apr 2010;5(4):673-682. [doi: [10.2215/CJN.07891109](https://doi.org/10.2215/CJN.07891109)] [Medline: [20338960](https://pubmed.ncbi.nlm.nih.gov/20338960/)]
46. Casagrande SS, Cowie CC, Fradkin JE. Utility of the U.S. Preventive Services Task Force criteria for diabetes screening. *Am J Prev Med*. Aug 2013;45(2):167-174. [doi: [10.1016/j.amepre.2013.02.026](https://doi.org/10.1016/j.amepre.2013.02.026)] [Medline: [23867023](https://pubmed.ncbi.nlm.nih.gov/23867023/)]
47. Young TK, Mustard CA. Undiagnosed diabetes: does it matter? *CMAJ*. Jan 9, 2001;164(1):24-28. [Medline: [11202663](https://pubmed.ncbi.nlm.nih.gov/11202663/)]
48. Tenenbaum A, Motro M, Fisman EZ, et al. Clinical impact of borderline and undiagnosed diabetes mellitus in patients with coronary artery disease. *Am J Cardiol*. Dec 15, 2000;86(12):1363-1366. [doi: [10.1016/s0002-9149\(00\)01244-3](https://doi.org/10.1016/s0002-9149(00)01244-3)] [Medline: [11113414](https://pubmed.ncbi.nlm.nih.gov/11113414/)]
49. Wang L, Li X, Wang Z, et al. Trends in prevalence of diabetes and control of risk factors in diabetes among US adults, 1999-2018. *JAMA*. Jun 25, 2021;326(8):1-13. [doi: [10.1001/jama.2021.9883](https://doi.org/10.1001/jama.2021.9883)] [Medline: [34170288](https://pubmed.ncbi.nlm.nih.gov/34170288/)]
50. Khan B, Fatima H, Qureshi A, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. *Biomed Mater Devices*. Feb 8, 2023;8 2023:1-8. [doi: [10.1007/s44174-023-00063-2](https://doi.org/10.1007/s44174-023-00063-2)] [Medline: [36785697](https://pubmed.ncbi.nlm.nih.gov/36785697/)]
51. Habebh H, Gohel S. Machine learning in healthcare. *Curr Genomics*. Dec 16, 2021;22(4):291-300. [doi: [10.2174/1389202922666210705124359](https://doi.org/10.2174/1389202922666210705124359)] [Medline: [35273459](https://pubmed.ncbi.nlm.nih.gov/35273459/)]

52. A Romero RA, Y Deypalan MN, Mehrotra S, et al. Benchmarking AutoML frameworks for disease prediction using medical claims. *BioData Min.* Jul 26, 2022;15(1):15. [doi: [10.1186/s13040-022-00300-2](https://doi.org/10.1186/s13040-022-00300-2)] [Medline: [35883154](https://pubmed.ncbi.nlm.nih.gov/35883154/)]
53. Rashidi HH, Tran N, Albahra S, Dang LT. Machine learning in health care and laboratory medicine: general overview of supervised learning and Auto-ML. *Int J Lab Hematol.* Jul 2021;43 Suppl 1:15-22. [doi: [10.1111/ijlh.13537](https://doi.org/10.1111/ijlh.13537)] [Medline: [34288435](https://pubmed.ncbi.nlm.nih.gov/34288435/)]
54. Raj R, Kannath SK, Mathew J, Sylaja PN. AutoML accurately predicts endovascular mechanical thrombectomy in acute large vessel ischemic stroke. *Front Neurol.* 2023;14:1259958. [doi: [10.3389/fneur.2023.1259958](https://doi.org/10.3389/fneur.2023.1259958)] [Medline: [37840939](https://pubmed.ncbi.nlm.nih.gov/37840939/)]
55. Yang J, Shi R, Wei D, et al. MedMNIST v2—a large-scale lightweight benchmark for 2D and 3D biomedical image classification. *Sci Data.* Jan 19, 2023;10(1):41. [doi: [10.1038/s41597-022-01721-8](https://doi.org/10.1038/s41597-022-01721-8)] [Medline: [36658144](https://pubmed.ncbi.nlm.nih.gov/36658144/)]
56. Zhuhadar LP, Lytras MD. The application of AutoML techniques in diabetes diagnosis: current approaches, performance, and future directions. *Sustainability.* 2023;15(18):13484. [doi: [10.3390/su151813484](https://doi.org/10.3390/su151813484)]
57. Mohsen F, Biswas MR, Ali H, Alam T, Househ M, Shah Z. Customized and automated machine learning-based models for diabetes type 2 classification. *Stud Health Technol Inform.* Jun 29, 2022;295:517-520. [doi: [10.3233/SHTI220779](https://doi.org/10.3233/SHTI220779)] [Medline: [35773925](https://pubmed.ncbi.nlm.nih.gov/35773925/)]
58. Khushi M, Shaikat K, Alam TM, et al. A comparative performance analysis of data resampling methods on imbalance medical data. *IEEE Access.* 2021;9:109960-109975. [doi: [10.1109/ACCESS.2021.3102399](https://doi.org/10.1109/ACCESS.2021.3102399)]
59. Echouffo-Tcheugui JB, Selvin E. Prediabetes and what it means: the epidemiological evidence. *Annu Rev Public Health.* Apr 1, 2021;42:59-77. [doi: [10.1146/annurev-publhealth-090419-102644](https://doi.org/10.1146/annurev-publhealth-090419-102644)] [Medline: [33355476](https://pubmed.ncbi.nlm.nih.gov/33355476/)]
60. Skyler JS, Bakris GL, Bonifacio E, et al. Differentiation of diabetes by pathophysiology, natural history, and prognosis. *Diabetes.* Feb 2017;66(2):241-255. [doi: [10.2337/db16-0806](https://doi.org/10.2337/db16-0806)] [Medline: [27980006](https://pubmed.ncbi.nlm.nih.gov/27980006/)]
61. Huguet N, Chen J, Parikh RB, et al. Applying machine learning techniques to implementation science. *Online J Public Health Inform.* Apr 22, 2024;16:e50201. [doi: [10.2196/50201](https://doi.org/10.2196/50201)] [Medline: [38648094](https://pubmed.ncbi.nlm.nih.gov/38648094/)]
62. Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci.* 2021;2(3):160. [doi: [10.1007/s42979-021-00592-x](https://doi.org/10.1007/s42979-021-00592-x)] [Medline: [33778771](https://pubmed.ncbi.nlm.nih.gov/33778771/)]
63. Shick AA, Webber CM, Kiarashi N, et al. Transparency of artificial intelligence/machine learning-enabled medical devices. *NPJ Digit Med.* Jan 26, 2024;7(1):21. [doi: [10.1038/s41746-023-00992-8](https://doi.org/10.1038/s41746-023-00992-8)] [Medline: [38273098](https://pubmed.ncbi.nlm.nih.gov/38273098/)]

Abbreviations

ADA: American Diabetes Association

AUC: area under the receiver operating characteristic curve

AutoML: automated machine learning

CREMLS: Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Models

MET: metabolic equivalent

NHANES: National Health and Nutrition Examination Survey

NPV: negative predictive value

PPV: positive predictive value

XGBoost: extreme gradient boosting

Edited by Khaled El Emam; peer-reviewed by Harsh Maheshwari, Jerry John Thayil; submitted 31.10.2024; final revised version received 23.05.2025; accepted 22.07.2025; published 08.10.2025

Please cite as:

Liu J, Ssewamala F, An R, Ji M

Use of Automated Machine Learning to Detect Undiagnosed Diabetes in US Adults: Development and Validation Study

JMIR AI 2025;4:e68260

URL: <https://ai.jmir.org/2025/1/e68260>

doi: [10.2196/68260](https://doi.org/10.2196/68260)

© Jianxiu Liu, Fred Ssewamala, Ruopeng An, Mengmeng Ji. Originally published in JMIR AI (<https://ai.jmir.org>), 08.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR AI, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.ai.jmir.org/>, as well as this copyright and license information must be included.