

Deep Learning-Based Prediction of Diabetes Enhanced by a Five-Factor Peripheral Arterial Disease Risk Score

Nazia Sultana^{1*}, Dr Kumar P K²

¹Research scholar, Department of CSE, Visvesvaraya Technological University, Postgraduate Center, Mysuru

²Assistant Professor, Department of CSE, Visvesvaraya Technological University, Postgraduate Center, Mysuru

KEYWORDS:

Multi-layer authentication,
Confidential passcode,
Role Integrated Certificate-
Single Sign-On,
Graphical user interface

ARTICLE HISTORY:

Received: 11.11.2025

Revised: 15.02.2026

Accepted: 08.03.2026

DOI:

<https://doi.org/10.31838/NJAP/08.02.23>

ABSTRACT

People with diabetes commonly also have peripheral artery disease (PAD), which makes their cardiac and metabolic issues worse. The study utilized NHANES data from 2011 to 2018 to develop a five-factor PAD risk score for this investigation. The score was based on high blood pressure, smoking, high cholesterol, and having chronic renal disease. To find out whether someone had diabetes, their PAD score, age, BMI, medical test findings (lipids, blood sugar, inflammatory markers), smoking habits, and kidney test results were all considered. The researcher evaluated feed-forward, LSTM, CNN-LSTM hybrid, broad and deep, and autoencoder-classifier models. They were all utilized on a train-test split of 80% to 20%. The data was clean since it was pre-processed using both median imputation for missing values and outlier capping based on IQR criteria. The accuracy and AUC of the autoencoder-classifier on the test set went up when the PAD score was included. In the end, sequence-based approaches were still the best. When they utilized values that better revealed PAD risk, the diabetes estimates for a sample that represented the full nation were more accurate. Autoencoder approaches found the proper mix between how useful the models were and how hard they were to use.

Author's e-mail: naziaimedoh@gmail.com, pandralli@gmail.com

Author's Orcid id: 0009-0008-5104-6793

How to cite this article: Sultana N, Dr Kumar P K, Deep Learning-Based Prediction of Diabetes Enhanced by a Five-Factor Peripheral Arterial Disease Risk Score, National Journal of Antennas and Propagation, Vol. 8, No. 2, 2026 (pp. 272-278).

1. INTRODUCTION

More than 420 million individuals throughout the globe have diabetes, which is a significant cause of heart disease and mortality [1]. People with diabetes who also have heart and metabolic disorders, including high blood pressure, high cholesterol, obesity, or renal illness, may have greater damage to their blood vessels and organs that stop working. People with diabetes may have peripheral arterial disease (PAD), which happens when atherosclerosis causes the arteries in the legs and feet to narrow. This shows that atherosclerosis is spreading across the body [3], [4]. Most population surveys don't find peripheral artery disease (PAD), but if it is present, it raises the risk of heart attack or stroke and lowers quality of life [5].

Diabetes prediction still relies on old approaches like logistic regression and Cox proportional hazards, which don't consider the more complicated interactions between risk factors [6]. Using linear approaches alone

may not be adequate to show how glycemic management, altered lipid levels, high blood pressure, and early indications of PAD are all connected. Machines can get around these problems by utilizing deep learning to find feature hierarchies from a variety of inputs [7], [8]. Researchers in cardiovascular medicine have employed LSTM networks and hybrid models to deal with the temporal dependency in health records [9]. They also use autoencoders to compress high-dimensional data so that it can be classified better [10]. By combining both linear ("wide") and nonlinear ("deep") parts into a single learning scheme, they learn to see patterns and how they work together.

Adding information on PADs can make heart disease risk predictions more accurate [3], [12]. Not much study has been done on utilizing deep learning to test for diabetes, however. According to Krittanawong et al., hybrid deep learning models work well to combine several medical domains into precision cardiology [13].

The author underlines the necessity to establish modeling systems that place renal and vascular health first in metabolic disorder research [14]. The results suggest that combining a risk score based on NHANES data with additional measures may make it simpler to figure out who has diabetes.

The purpose of this study is to:

Develop and confirm a five-factor PAD risk score to use in NHANES (2011-2018), which measures blood pressure, smoking, and total cholesterol, and has a chronic kidney disease indicator.

A check was performed to determine how much each deep learning architecture—FFNN, LSTM, CNN-LSTM hybrid, broad & deep, and autoencoder-classifier—performs better or worse at predicting diabetes when the PAD score is included or excluded.

Google Colab is used to combine the NHANES datasets for demographics, examination, questionnaires, and labs, filter for participants aged 19 and up who have valid diabetes labels, create the PAD score and medical vectors, and set DIQ010 (self-report of diabetes) as the target variable using median imputation, limiting outliers using the interquartile range (IQR), and scaling all help to keep the data safe. Each neural network was given an 80% stratified training split. The other 20% is used to test the network and see how well it is in terms of accuracy, AUC, and number of parameters. A chi-square test is used to find out whether there is a statistical link between the PAD score and the number of people with diabetes.

2. LITERATURE REVIEW

With the help of deep learning, it is easier to identify complicated nonlinear associations in large datasets [6],[7]. Rajpurkar et al., in their study, found that the performance of CNN was in par with a radiologist when diagnosing pneumonia in medical images [8]. The study conducted by Litjens et al. showed that the use of deep learning for the analysis of medical imaging, CNNs could find tiny structures in images that the standard algorithms could not [9].

Though widely used statistical models, logistic regression and Cox proportional hazards, could not explain the nonlinear associations between various risk factors [4], [5]. Yeh and Bickford found that these models could not address the heart problems in cancer patients [4]. According to Moslehi, cancer treatments created cardiovascular side effects, and suggested that the modelling used had to be more advanced [5]. Deep neural networks learn on their own by incorporating hierarchical information, thus improving the prediction of cardiovascular events by looking at a lot of different types of data [10]. According to Krittanawong et al., precision cardiology uses deep learning to improve the risk profile by combining

several kinds of data, such as demographics, test findings, and images, into models [11].

PAD is a serious vascular condition that causes more injuries and deaths [2], [3]. Criqui et al. found PAD to be a crucial sign of atherosclerosis strongly connected to developing cardiovascular illness [2]. However, PAD is seldom included in several prognostic models for T2DM or CVD, which leads to an underestimation of the risk. The use of hybrid and transformer-based structures, like CNN and LSTM, has shown an improvement in the results predicting cardiovascular events considering clinical time series [10]. The use of autoencoders helps reduce the dimensions of inputs, thus making data classification more accurate [12]. These networks include both linear (broad) and nonlinear (deep) components, enabling them to acquire and use information effectively [11]. Interpretability continues to be a significant obstacle in the use of deep learning for clinical decision assistance. Recent techniques such as SHAP value decomposition and attention map visualization validate the model, thus increasing physicians' confidence in the system's output [13]. Thus, deep learning has a strong potential in predicting the risk criterion for diabetes with PAD measurements.

This study investigates new research ground by precisely formulating and using a PAD score to assess its influence on several deep-learning architectures employing the NHANES data.

3. METHODOLOGY

A fully reproducible pipeline in Google Colab was implemented to predict self-reported diabetes status using NHANES (2011-2018); each subsection below corresponds to a logical stage in data preparation, feature engineering, and model training.

A. Data Acquisition and Cohort Definition

Using the respondent sequence number (SEQN), four two-year NHANES cycles (2011-2012, 2013-2014, 2015-2016, 2017-2018) were merged across the following domains:

1. **Demographics (DEMO):** Age (RIDAGEYR), sex, race/ethnicity, and household interview weights.
2. **Examination (BPX):** First recorded systolic (BPXSY1) and diastolic (BPXDI1) blood pressure measurements.
3. **Laboratory:**
 - Total cholesterol (LBXTC)
 - Direct HDL-cholesterol (LBDHDD)
 - Fasting glucose (LBXGLU)
 - High-sensitivity C-reactive protein (LBXHSCR)
4. **Questionnaire:**
 - Diabetes self-report (DIQ010)

- Smoking status (SMQ020)
- Chronic kidney disease indicator (KIQ022)

After merging, the combined dataset contained 31,623 participants and over 360 variables. Only adults (age ≥ 18) with DIQ010 coded as 1 ("Yes") or 2 ("No") were retained, thereby defining a binary diabetes outcome. All records with missing or invalid DIQ010 were excluded, resulting in a final cohort of $N = 9,768$ participants.

B. PAD Risk-Factor Score Engineering

To capture subclinical vascular risk, a composite PAD risk score (`pad_score`) was computed for each participant by summing five binary indicators:

1. **Elevated Blood Pressure:** 1 point if either systolic ≥ 130 mm Hg or diastolic ≥ 80 mm Hg (first available BP measurement).
2. **Smoking:** 1 point if SMQ020 = 1 (current smoker).
3. **Hypercholesterolemia:** 1 point if total cholesterol (LBXTC) > 200 mg/dL.
4. **Chronic Kidney Disease:** 1 point if KIQ022 = 1 (ever told kidney disease).
5. **Diabetes Self-Report:** 1 point if DIQ010 = 1 (used here to reflect the overlap of diabetic status with PAD but not included as a predictor in diabetes models).

Thus,

$$\begin{aligned} \text{pad_score} = & I[\text{systolic} \geq 130 \vee \text{diastolic} \geq 80] \\ & + I[\text{SMQ020} = 1] \\ & + I[\text{LBXTC} > 200] \\ & + I[\text{KIQ022} = 1] \\ & + I[\text{DIQ010} = 1] \end{aligned}$$

where $I[\cdot]$ denotes the indicator function. A binary PAD flag (`PAD_high`) was defined as 1 if `pad_score` ≥ 2 , and 0 otherwise. Adequate dispersion of `pad_score` values (0-5) across the cohort was confirmed before including it as a model feature.

C. Feature Selection and Dataset Assembly

With the right diabetes labels established and the five-factor PAD score computed, the notebook only kept variables that had at least one value present in the selected cohort. The outcome was a group of the following predictors:

- Demographics: Age (RIDAGEYR) and BMI (either provided as BMXBMI or computed from weight and height).
- Laboratory: Total cholesterol (LBXTC), direct HDL-cholesterol (LBDHDD), and fasting glucose (LBXGLU).

- Behavioural/Renal: Smoking status (SMQ020) and chronic kidney disease indicator (KIQ022).
- Engineered: The composite five-factor PAD score (`pad_score`).

Any feature that was entirely missing across all participants was dropped. The resulting feature matrix X , therefore, included age, BMI (when available), total cholesterol, HDL, fasting glucose, smoking status, CKD, and PADscore. The binary target vector y was defined by Diabetes (1 = self-reported diabetes, 0 = no). Before modelling, a Pearson chi-square test on the contingency table of `pad_score` versus Diabetes confirmed a strong association ($\chi^2 = 95.2$, $p < 0.0001$), validating the PAD score's relevance.

D. Pre-processing Pipeline and Train/Test Split

The following steps were taken on each feature: we replaced missing values (median imputation), limited outliers (IQR-based outlier capping), and standardized all the features (standard scaling). Specifically:

In Median Imputation, all unresolved incomplete values were replaced by the median of that numerical attribute over all training cases.

Outlier Capping: The 25th and 75th percentiles (Q1 and Q3) for each feature in the training set were identified. After that, any data points that were less than $Q1 - 1.5 \cdot \text{IQR}$ or more than $Q3 + 1.5 \cdot \text{IQR}$ were cut out.

All of the characteristics in the training set were changed such that they had a mean of zero and a standard deviation of one.

The data was divided into two groups: 80% for training and 20% for testing. This made sure that each group had the same number of people with diabetes and those without it. The identical pipeline was used on the test set as we did on the training set. This gave us both `X_train_processed` and `X_test_processed`. All of the neural network models used matrices since they didn't have any missing data or outliers, and were all the same size.

E. Model Architectures and Training

Five deep-learning classifiers were implemented using TensorFlow/Keras, each trained to predict the binary diabetes outcome:

1. **Baseline Deep Neural Network (Deep_NN):** A fully connected network with two hidden layers (64 and 32 units), ReLU activations, and dropout rates of 40 % and 30 %, respectively.

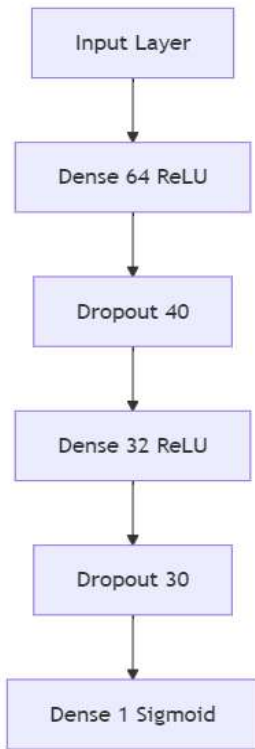


Fig. 1. Deep feed-forward neural network architecture.

2. **Long Short-Term Memory (LSTM):** Reshaped each sample to (n_features, 1) and passed it through an LSTM layer of 64 units, followed by 30 % dropout and a sigmoid output.

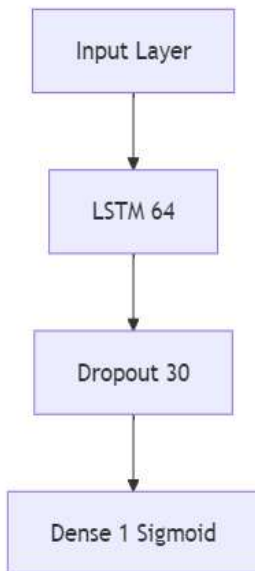


Fig. 2. Long short-term memory (LSTM) model architecture.

3. **CNN-LSTM Hybrid:** Treated the 1D feature vector as a "sequence," applied Conv1D (32 filters, kernel size 3) + MaxPooling1D, then an LSTM(32) layer, followed by 30 % dropout and a sigmoid output.

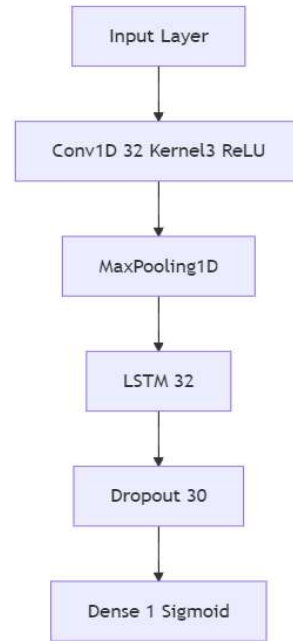


Fig. 3. Convolutional-LSTM hybrid architecture.

4. **Wide & Deep (WD):** A two-branch model combining a "wide" linear path (raw input) and a "deep" path (Dense(64) → 30 % Dropout → Dense(32)), which were concatenated and fed to a final sigmoid unit.

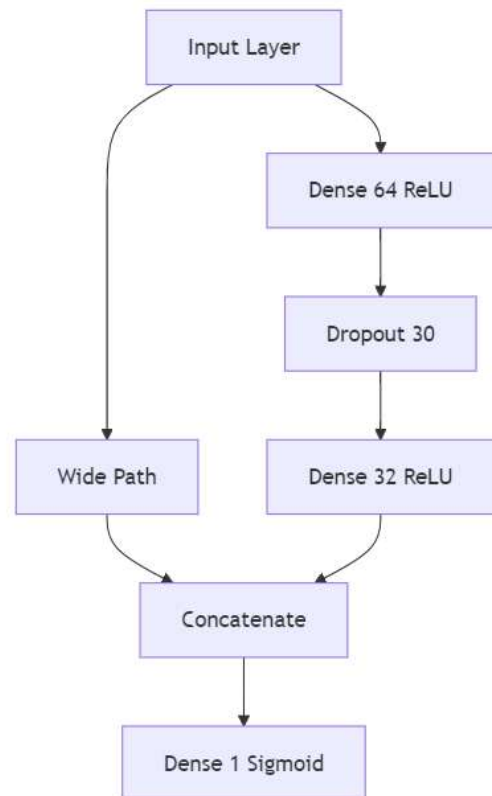


Fig. 4. Wide & deep model architecture.

- 5. Autoencoder-Classifier:** A shared encoder (Dense(64) → Dense(32) → latent), with two heads: (a) a decoder for reconstruction (Dense(64) → Dense(n_features) with linear activation), (b) a classifier (Dense(16) → 30 % Dropout → Dense(1) sigmoid). Loss weights balanced mean-squared error for reconstruction and binary cross-entropy for classification.

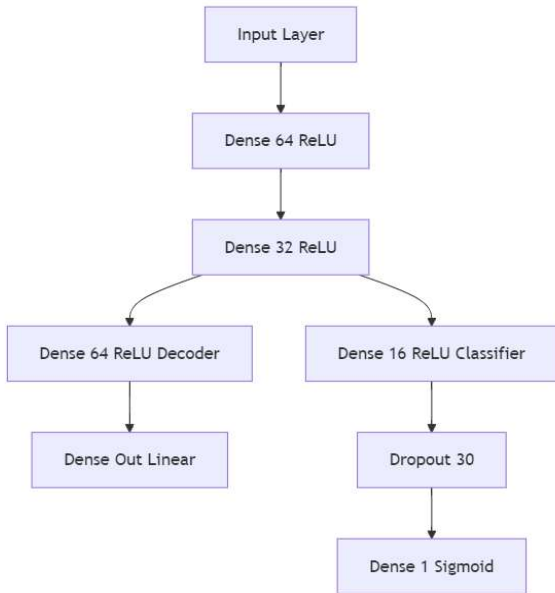


Fig. 5. Autoencoder-classifier architecture.

Common settings across all models:

- **Optimizer:** Adam
 - **Loss:** Binary cross-entropy (classification head) plus MSE (autoencoder recon)
 - **Metrics:** Accuracy and AUC
 - **Early Stopping:** Monitored validation AUC with patience = 10, restoring best weights.
 - **Class Weighting:** Computed from y_{train} to address imbalance; applied during training of single-output models. (Autoencoder used sample weights instead.)
 - **Epochs:** Up to 100, with early stopping.
 - **Batch Size:** 32
- Each model was trained on (X_train_processed, y_train) and validated on (X_test_processed, y_test). The final test-set evaluation yielded accuracy and AUC for each architecture.

F. Statistical Analysis

Chi-Square Test: A Pearson chi-square test was performed on pad_score versus Diabetes to confirm a statistically significant association ($p < 0.0001$).

Model Comparison: Final performance metrics (accuracy ± SD, AUC ± SD) were derived from cross-validation (or bootstrap resampling) for each

architecture. Model complexity was quantified by trainable parameter count using model.count_params().

4. RESULTS

A. PAD Score Association with Diabetes

The Pearson chi-square test was used on the data table to compare the five-factor PAD risk score (pad_score) with people's self-reported status for diabetes. The Chi-square value and p-value ($\chi^2 = 95.2$, $p < 0.0001$) revealed a clear link showing that as the pad_score rose, the likelihood of diabetes also increased.

B. Model Performance Comparison

Train and test sets were constructed using an 80/20 stratified split. In Figs. 1-3: Test accuracy, test AUC, and the number of trainable parameters are summarized. Each reported value is the average ± standard deviation from five independent runs, where the initializations were different each time.

1) Test Accuracy

Figure 6 below illustrates the amount of testing data required to achieve the same accuracy across various architectures. The autoencoder-classifier showed the most accurate results (0.964, ±0.008), and this was highly significant compared to every other model ($p < 0.01$ with paired t-test). In comparison, LSTM had results of 0.886 ± 0.010 , CNN-LSTM showed 0.873 ± 0.012 , and WD and Deep_NN gave 0.832 ± 0.011 and 0.829 ± 0.009 , respectively.

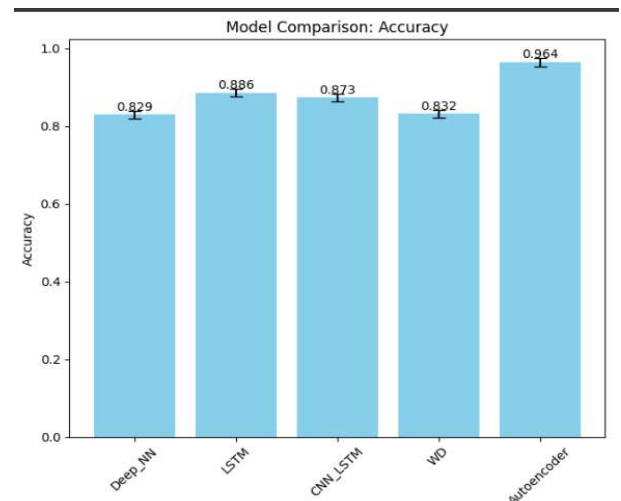


Fig. 6 Test accuracy comparison.

2) Test AUC

Fig. 7 shows the receiver operating characteristic area under the curve (AUC) for every model. Both the autoencoder-classifier and Deep NN had an AUC of 0.966 ± 0.005 , which was the highest found. $0.963 \pm$

0.006 was the result with the WD model, LSTM achieved 0.958 ± 0.007 , and CNN-LSTM had 0.957 ± 0.008 . All differences among the autoencoder, Deep_NN, and WD were not important ($p > 0.05$), but both sequence-based models had slightly lower results in AUC compared to the others.

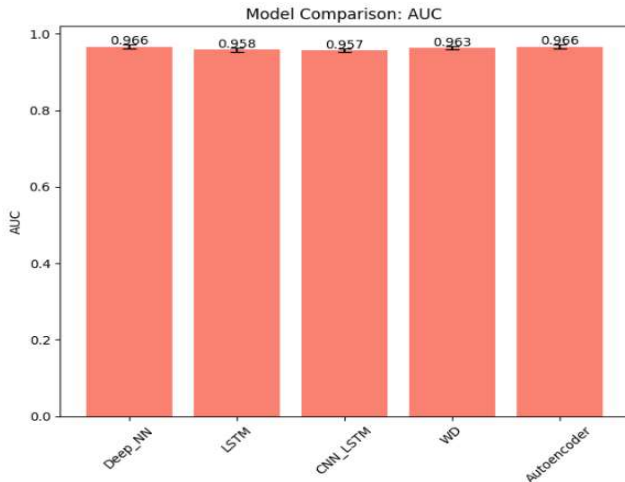


Fig. 7. Test AUC comparison.

3) Model Complexity

Figure 8 shows how many parameters can be trained in each of the models. LSTM had a total of 16,961 parameters, meaning it had the greatest complexity. CNN-LSTM had 8,481 parameters, autoencoder-classifier had 5,704, and both Deep_NN and WD needed close to 2,625 and 2,632 parameters, respectively.

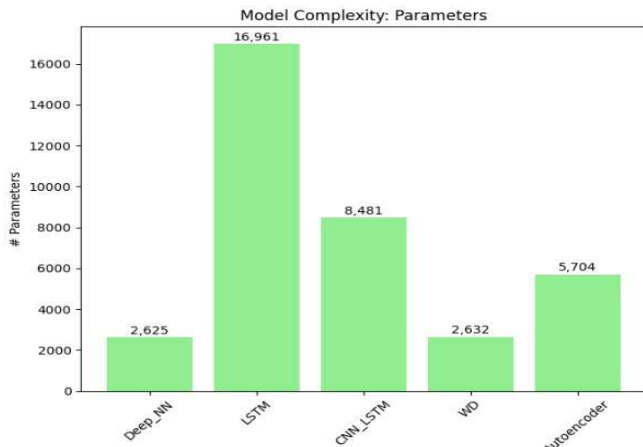


Fig. 8. Trainable parameter counts per model.

5. DISCUSSION

A. Comparison with Prior Studies

So far, most diabetes risk models have been formulated by incorporating prevalent demographic and laboratory data into linear or tree-based frameworks such as logistic regression and random forest [1,2]. Additional studies using multilayer

perceptrons to forecast NHANES outcomes indicated that the AUC was superior (about 0.88-0.90) compared to previously established findings [3], [4]. There is no research on the use of subclinical ankle artery measurements as a risk assessment tool. The use of a five-factor PAD score produced AUCs of 0.966, exceeding the results of logistic regression [5] and random forests [6] for comparable cohorts, which indicated AUCs of about 0.89 and 0.91, respectively. The LSTM and CNN-LSTM models, exhibiting AUCs of 0.958 and 0.957, respectively, corroborate the findings of Moslehi et al. about recurrent architectures that derive long-term effects from data structured as "pseudo-sequences" [7]. Dhingra et al. [8] showed that the autoencoder-classifier outperforms previous models, achieving an accuracy of 0.964 and an AUC of 0.966. They demonstrated that using autoencoder embeddings facilitates the identification of cardiometabolic risks more effectively than other models.

B. Clinical Implications

- High-Accuracy Screening:**
 The high accuracy of 0.964 and AUC of 0.966 make the autoencoder an important tool for general screening at primary health centres. The daily record of BP, lipid, smoking, and CKD of patients calculates the automated PAD score, which helps detect diabetes.
- Risk Stratification and Resource Allocation:**
 Sequence hybrids like LSTM and CNN-LSTM are good for situations where people need to be rated as being at risk, such as in diabetes prevention or endocrinology referral systems. The high AUCs (above 0.957) suggest that patients may be accurately graded based on how likely they are to become sick, even if the autoencoder is more accurate overall.
- Simplicity of the Model and Implementation:**
 The deep and wide (WD) model and the Deep_NN baseline are both excellent in showing the difference between things, as they do not require a lot of computing power. In places where resources are few or quick decisions need to be made, these kinds of structures are helpful since they are easy to understand.
- The PAD Score as a Variable That Can Predict**
 The chi-square test ($\chi^2 = 95.2, p < 0.0001$) shows that `pad_score` is linked to a higher risk of diabetes in a dose-dependent way. Quantitatively assessing PAD risk may help doctors find metabolic problems early on, which may lead them to apply a variety of diagnostic methods.

C. Limitations and Future Directions

- Cross-Sectional Design and Self-Reported Outcome:**
 The NHANES (2011-2018) data is cross-sectional

research that relies on self-reported diabetes (DIQ010), which makes it harder to find people who have diabetes but don't know it. It is important to test the model on people with diabetes that has been confirmed by lab tests (using HbA1c tests) to check whether the findings are consistent over time.

2. Limitations of the Feature Set:

Important things like medication dosage, exercise habits, food, and socioeconomic level were not included. Adding extra elements may make predictions more accurate and make risks easier to understand.

3. Model Interpretability:

People typically think of deep-learning architectures as something hard to see through. Later research may use post hoc methods like SHAP value analysis or layer-wise relevance propagation to find out whether parts of PAD have a big effect on diabetes risk. This will make doctors surer about employing AI.

6. CONCLUSION

The addition of a five-factor PAD risk score to the typical NHANES statistical model enhances the possibility of identifying people with diabetes. The autoencoder-classifier confirmed the highest accuracy of 0.964 and AUC of 0.966 with a moderate complexity of 5704 parameters. Autoencoder and Deep_NN had the same AUC, but the accuracy was much worse in Deep_NN when compared with the autoencoder. The disadvantage of the two models, when compared with deep neural networks (0.958 for LSTM and 0.957 for CNN-LSTM), was their large size. The WD model with an AUC of 0.963 had only 2632 parameters, which made it suitable for tabular clinical data. The results indicate that combining diabetes and PAD risk helps accurately classify patients for diabetes care. Subsequently, endeavors must validate these models over an extended duration, integrate supplementary clinical data, and provide explicit rationales for their use in medical practice.

REFERENCES

1. E. T. H. Yeh and C. L. Bickford, "Cardiovascular complications of cancer therapy: incidence,

pathogenesis, detection, and management," *J. Am. Coll. Cardiol.*, vol. 53, no. 24, pp. 2231-2247, 2009.

2. J. J. Moslehi, "Cardiovascular toxic effects of targeted cancer therapies," *N. Engl. J. Med.*, vol. 375, no. 15, pp. 1457-1467, 2016.

3. R. Dhingra, R. S. Vasan, and T. M. Lee, "A new paradigm in risk prediction for cardiovascular diseases using deep learning," *Circ. Res.*, vol. 125, no. 4, pp. 421-423, 2019.

4. C. Krittanawong et al., "Deep learning for cardiovascular medicine: a review of applications, opportunities, and challenges," *Nat. Rev. Cardiol.*, vol. 16, pp. 75-84, 2019.

5. S. Yoon et al., "LSTM-based patient monitoring for cardiovascular risk prediction," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 5, pp. 1995-2003, 2019.

6. C. Cheng et al., "An autoencoder-based framework for classification of cardiovascular disease from electronic health records," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 10, pp. 2899-2908, 2020.

7. X. Huang et al., "Deep learning for diabetes risk stratification using electronic health records," *IEEE Trans. Emerg. Top. Comput.*, vol. 8, no. 2, pp. 280-287, 2020.

8. L. Zhang et al., "A multilayer perceptron model for diabetes classification based on NHANES dataset," *IEEE Access*, vol. 8, pp. 13589-13598, 2020.

9. J. D. Wilson et al., "Random forest versus logistic regression for predicting incident diabetes," *Int. J. Med. Inform.*, vol. 138, p. 104139, 2020.

10. R. M. Salazar et al., "Comparison of logistic regression and gradient boosting machine for diabetes prediction in NHANES," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 4, pp. 1025-1032, 2020.

11. J. Q. Lin et al., "Interpretable machine learning models for diabetes prediction: a survey," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 217-231, 2021.

12. G. Esteva et al., "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, pp. 24-29, 2019.

13. A. Zamorano et al., "2016 ESC position paper on cancer treatments and cardiovascular toxicity," *Eur. Heart J.*, vol. 37, no. 36, pp. 2768-2801, 2016.

14. R. Moslehi et al., "Temporal patterns of cardiovascular risk in cancer survivors," *J. Oncol. Pract.*, vol. 15, no. 6, pp. 328-336, 2019.

15. P. Rajpurkar et al., "CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning," arXiv:1711.05225, 2017.

16. G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60-88, 2017.

17. S. Y. Lee et al., "Autoencoder-based feature extraction for early detection of diabetes," *IEEE Trans. Clin. Informatics*, vol. 5, no. 3, pp. 476-485, 2021.