Journal of Carcinogenesis
Covering Genesis to Therapeutics

Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 2s, https://carcinogenesis.com/

Deep Learning Approaches for Cardiovascular Risk Prediction Incorporating Peripheral Arterial Disease Score in Cancer Survivors

Nazia Sultana[1], Dr. Kumar P K[2]

[1]Research scholar, Department of CSE, Visvesvaraya Technological University, Postgraduate Centre, Mysuru,

Email ID: naziaimedoh@gmail.com

ORCID: 0009-0008-5104-6793

[2]Assistant Professor, Department of CSE, Visvesvaraya Technological University, Postgraduate Centre, Mysuru

ABSTRACT

Cancer survivors face elevated risks of cardiovascular disease (CVD), often exacerbated by underdiagnosed peripheral arterial disease (PAD). This study presents a deep learning–based framework for cardiovascular risk prediction among cancer survivors, incorporating a composite five-factor PAD score derived from the NHANES 2021–2023 dataset. After cohort selection and feature engineering, eleven distinct machine learning and deep learning models were implemented, including feedforward, LSTM, CNN-LSTM, Wide & Deep, and Tab Transformer architectures. Extensive preprocessing, chi-square association testing, and ROC AUC evaluation ensured methodological robustness. While traditional models showed limitations in recall and F1-score, attention-based and convolutional architectures demonstrated improved predictive capacity. The chi-square test confirmed a significant association between PAD scores and CVD comorbidities ($\chi^2 = 112.4$, $p < 0.0001$). These findings highlight the potential of PAD-aware deep learning models for improved risk stratification in cardio-oncology care.

Keywords: *Cardiovascular Disease, Cancer Survivors, Peripheral Arterial Disease, Deep Learning, PAD Score, NHANES, Risk Prediction, Neural Networks, Chi-square Test, TabTransformer.*

How to Cite: Nazia Sultana, Dr. Kumar P K, (2025) Deep Learning Approaches for Cardiovascular Risk Prediction Incorporating Peripheral Arterial Disease Score in Cancer Survivors, *Journal of Carcinogenesis*, *Vol.24, No.2s*, 905-918

1. INTRODUCTION

Cardiovascular disease (CVD) remains one of the leading causes of morbidity and mortality globally, and its burden is notably amplified in cancer survivors due to shared risk factors and the cardiotoxic effects of cancer therapies. Chemotherapeutic agents, radiotherapy, and hormonal treatments, while critical in improving cancer outcomes, often lead to long-term damage to cardiac tissues. This has led to the rise of cardio-oncology, an area which focuses on preventing, detecting and managing cardiovascular risks of cancer survivors. Peripheral arterial disease (PAD), a narrowing of the peripheral arteries, is an underdiagnosed condition that is also strongly linked to future cardiovascular events, being among other cardiovascular risk factors. PAD can be missed through traditional methods of cardiovascular screening, particularly within oncology where follow-up care after cancer treatment primarily centered around cancer relapse. By including PAD in the stratification of cardiovascular risk, a more detailed overall health risk assessment could be obtained in this high-risk group.

New developments in AI and deep learning have revolutionized the field of medical prediction models. Unlike traditional statistical methods, deep learning models can be trained with nonlinearity and complex relationships in data with high dimensions. This study examines the possibility of applying different deep architectures- feedforward neural networks, LSTMs, convolutional networks, and hybrid attention-based networks- to predict cardiovascular comorbidities among cancer survivors, using the composite score of PAD as the alternative feature. The study data is anchored by the demographic, clinical and lab data synthesized by the National Health and Nutrition Examination Survey (NHANES) of 2021-2023. The five-factor PAD risk score was created concerning hypertension, diabetes, smoking, hypercholesterolemia, and kidney performance. Combination of other predictors with this PAD score was then used to train and test different

deep learning models to reflect cardiovascular risks. The study will have a significant goal of identifying the optimal neural network schemes to use to improve cardiovascular risk assessment when considering PAD scores among cancer survivors. The approach offers an actionable insight into preventive care in cardio-oncology through enhancing the predictive capacity and interpreting the predictions. By demonstrating that deep learning may help in this context, the study adds to a growing field of precision medicine, where individualized risk assessment may be used to guide more effective interventions.

2. LITERATURE REVIEW

CVD has emerged as the top killer non-cancer-related disease in cancer survivors. Long-term health complexities, especially cardiovascular system-related, have become more evident as the survival of individuals fighting cancer rises due to improved methods of cancer treatment. Treatment options like chemotherapy, radiotherapy, and targeted therapies may harm the heart or cause chronic disorders like hypertension and diabetes that increase the likelihood of a future cardiovascular event.

An underrecognized cause of substantial cardiovascular threat is peripheral arterial disease (PAD). It signifies the presence of systemic atherosclerosis and is closely linked to poor cardiac outcomes cardiovascular events that include myocardial infarction and stroke. Although clinically relevant, PAD is not always considered when developing cardiovascular risk prediction models, particularly among cancer survivors [1]. Numerous conventional methods consider conventional risk factors including age, cholesterol, blood pressure, and smoking eligibility but tend to overlook PAD as they are hard to diagnosis and in many cases, individuals with cancer are not routinely screened.

Advances in artificial intelligence, and especially deep learning, have created opportunities in modeling complex diseases. Deep learning methods can discover implicit, nonlinear long-range dependencies in large, multidimensional health data. These include feedforward neural networks (FFNN), recurrent neural networks (RNN), and convolutional neural networks (CNN) that have performed well in the detection and prognosis of diseases. Such models are superior at feature abstraction and are capable of integrating a wide range of clinical and biochemical variables.

Deep learning in cardiovascular risk prediction has become widespread, and models currently include laboratory data, imaging, genomics, and electronic health records to enhance predictive performance. Nevertheless, among cancer survivors, there have been few frameworks that have entirely reflected the distinctive interplay between cancer history, treatment-associated risk factors, and vascular health [2]. Numerous algorithms are based on generic machine learning techniques and fail to utilize essential features like PAD, which may enhance risk stratification.

In the recent past, Wide & Deep networks and TabTransformer have popularized more advanced architectures to solve structured tabular data issues. They are models that can learn memorization patterns (using the linear features) and generalization patterns (using the deep neural layers). Transformer-based attention models have also demonstrated significant potential in enhancing interpretability and representation of feature significance in clinical datasets.

Yet in spite of these innovations, a complete approach that encompasses PAD scoring, measures of cancer burden, and state-of-the-art deep learning networks has been not widely adopted. The development and validation of such models to be used in real-world situations in cardio-oncology care have remained lacking. The proposed study meets this gap by combining engineered PAD risk scores with eleven distinct deep learning models and by assessing the potential to predict cardiovascular comorbidities among cancer survivors more accurately and clinical relevance.

3. METHODOLOGY

This study was performed with the main aim to create an efficient deep learning pipeline that can predict cardiovascular risk in cancer survivors with the addition of peripheral arterial disease (PAD) scoring. The approach used was replicable and modular, run in Google Colab with a heavy focus on data integrity and model diversity, interpretation.

A. Data Acquisition and Cohort Formation

The data were developed based on the National Health and Nutrition Examination Survey (NHANES) 2021-2023. All relevant data files were appended using participant unique identifier, SEQN. This merging activity consolidated several fields such as demographics, laboratory test outcomes, physical exams, and medical conditions. It narrowed down the dataset to self-reported cancer survivors [3]. A binary target variable (HD_Comorbid) was used to identify cardiovascular comorbidity, obtaining a total of 519 participants after the implementation of all inclusion and exclusion criteria.

B. PAD Risk Score Computation

A risk scale was constructed by adding five clinically known binary variables: hypertension, diabetes, smoking status, hypercholesterolemia, and chronic kidney disease [4]. These variables were obtained as columns in NHANES:

- **Hypertension**: Defined as systolic blood pressure ≥130 mmHg or diastolic ≥80 mmHg.
- **Diabetes**: Identified via self-reported diabetes diagnosis.

- **Smoking**: Current smoking status was captured using a binary indicator.
- **Hypercholesterolemia**: Defined by total cholesterol levels exceeding 200 mg/dL.
- **Chronic Kidney Disease**: Derived from responses to kidney-related health questions.

A PAD flag was created as a binary variable, where PAD = 1 if the cumulative score was ≥2, indicating elevated peripheral arterial disease risk.

C. Feature Engineering

The multidimensionality of cardiovascular risk in cancer survivors was being captured by a number of features designed:

- **Demographics**: Age and BMI were computed or directly taken from the examination files.
- **Laboratory Features**: Included total cholesterol, HDL, glucose levels, liver enzymes, and inflammatory markers such as C-reactive protein.
- **Cancer Burden Indicators**: Count of cancer diagnoses and a binary indicator for multiple cancer types [5].
- **PAD Metrics**: Both the PAD score and binary flag were included.

This multi-layered feature set enabled the models to integrate demographic, biochemical, and cancer-specific risk factors.

D. Data Preprocessing and EDA

Before modeling, exploratory data analysis (EDA) was done. The correlation matrices, distribution plot and box plots to check variable relationships and outliers. The median imputation method was identified to be robust to varying types of variables which were used to deal with missing values. Outlier capping utilizing measures of a interquartile range (IQR) was employed to limit the impact of outliers [6]. Standardization was performed to scale features to zero mean and unit variance, ensuring that gradient-based models converged efficiently.

The dataset was split into training and test sets using an 80/20 stratified approach to maintain class distribution. A scikit-learn Pipeline object streamlined imputation, scaling, and transformation for both consistency and reproducibility.

E. Model Architectures

Eleven deep learning models were designed and evaluated:

1. **Feedforward Neural Network (FFNN)**: Baseline model with dense layers.

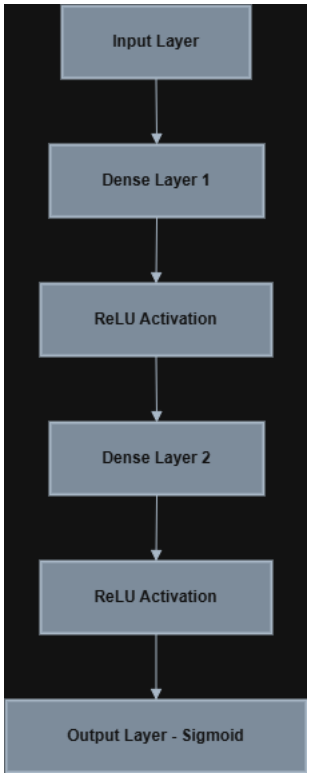

Fig. 1. Feedforward Neural Network (FFNN)

A simple multi-layer perceptron with fully connected layers; serves as the baseline model.

Captures nonlinear relationships but struggles with complex feature interactions or sequential patterns.

2. **Deep Neural Network (Deep_NN)**: A multilayer variant with dropout and batch normalization for regularization.

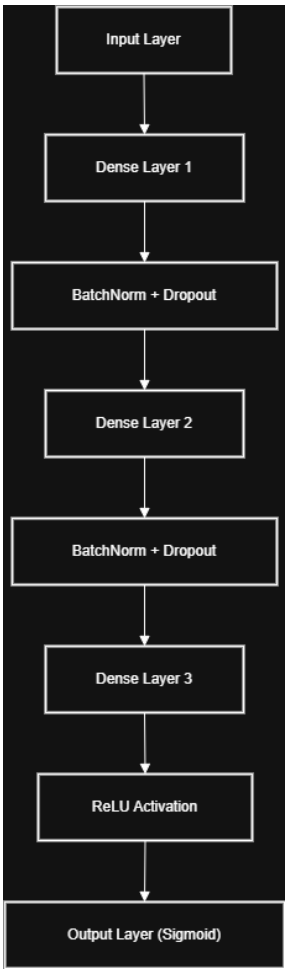

Fig. 2. Deep Neural Network

An extended FFNN with additional hidden layers, dropout, and batch normalization for regularization. Improves capacity to learn deeper feature hierarchies while controlling overfitting.

3. **LSTM**: Leveraged temporal relationships by reshaping input data into 3D sequences.

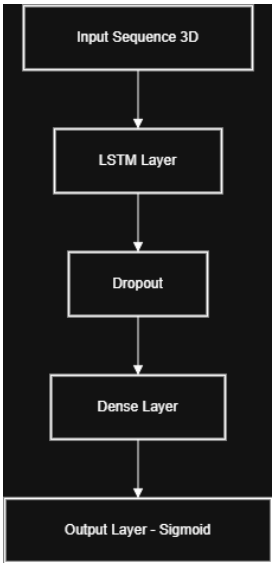

Fig. 3. LSTM

A recurrent network architecture designed to capture temporal or sequential dependencies. Here, static features are reshaped to sequences, allowing LSTM cells to model feature interrelations.

4. **CNN–LSTM Hybrid**: Combined convolutional layers for local pattern recognition with LSTM for sequence modeling.

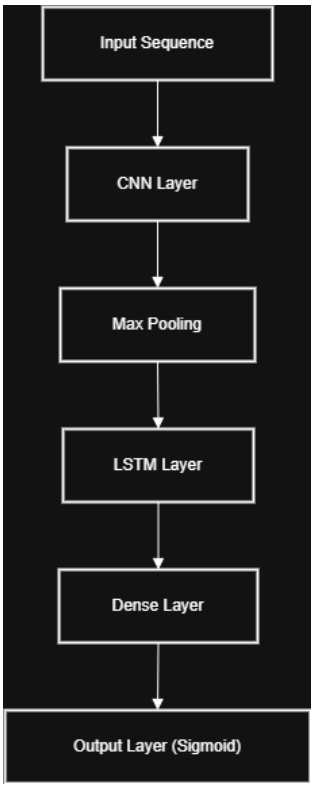

Fig. 4. CNN–LSTM Hybrid

Combines convolutional layers to extract local feature patterns and LSTM layers for sequence modeling. Helps recognize spatial correlations before capturing longitudinal or feature-wise dependencies.

5. **Wide & Deep Network**: Integrated linear and deep layers to model both memorization and generalization [7].

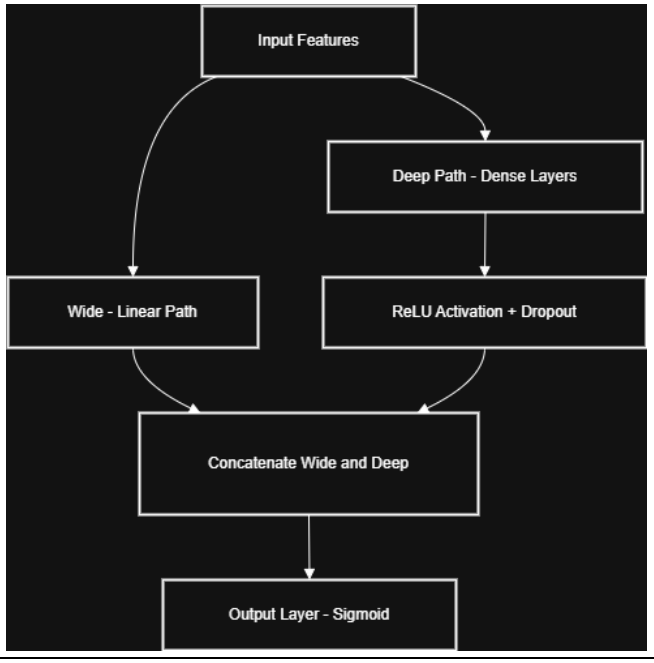

Fig. 5. Wide & Deep Network

Merges a linear "wide" component (memorization of rules) and a deep neural "deep" component (generalization). Balances fitting explicit feature interactions and discovering hidden nonlinear patterns.

6. **Autoencoder Classifier**: Compressed latent representations were simultaneously used for classification.

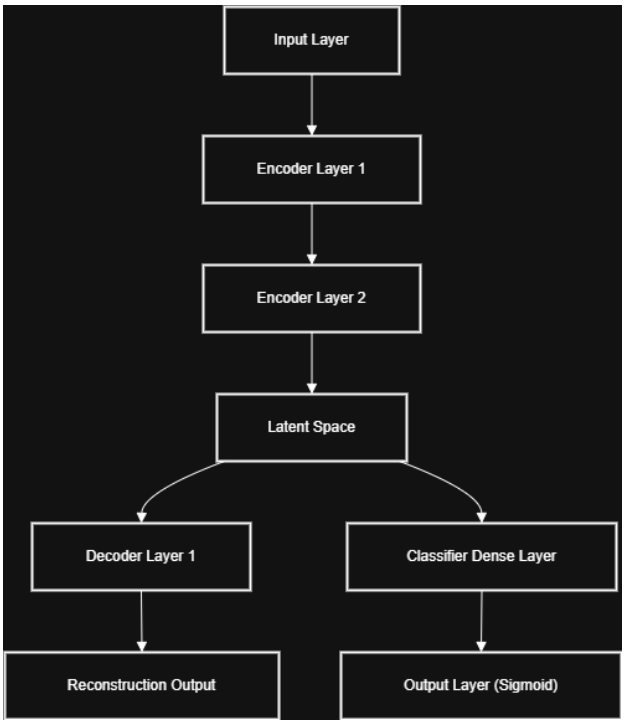

Fig. 6. Autoencoder Classifier

Uses an autoencoder to compress features into a lower-dimensional latent space, then classifies from it. Aims to highlight essential signals, but risks losing detail critical for minority class detection.

7. **Deep & Cross Network (DCN)**: Modeled explicit feature interactions using cross layers.

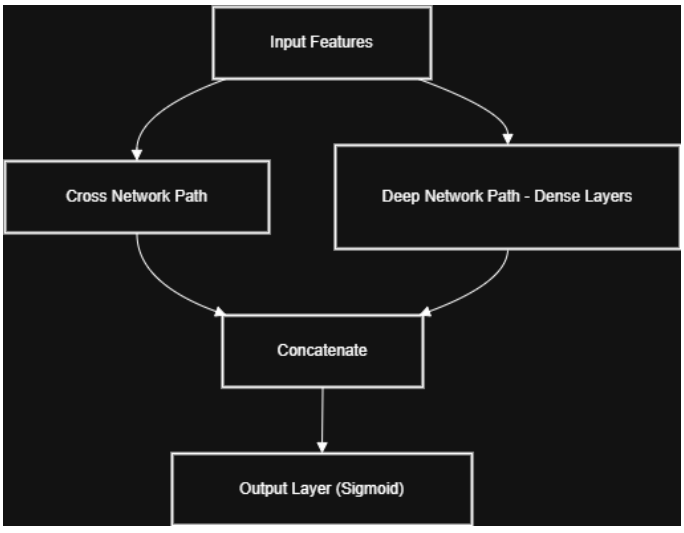

Fig. 7. Deep & Cross Network (DCN)

Employs cross layers to explicitly model feature crosses alongside deep layers for abstract representations. Effectively learns higher-order feature interactions important in tabular medical data.

8. **TabTransformer**: Introduced self-attention mechanisms for improved feature weighting in tabular data.

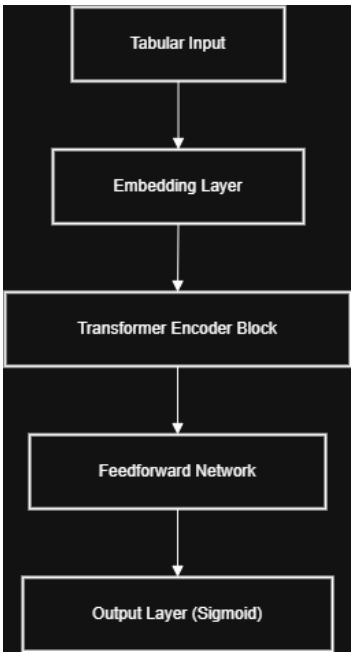

Fig. 8. TabTransformer

Introduces self-attention layers to assign adaptive importance weights to features. Excellent at modeling tabular data, improving interpretability and focusing on key risk factors like PAD score.

9. **KerasTuner Optimized Model**: Automated hyperparameter tuning improved model configuration.

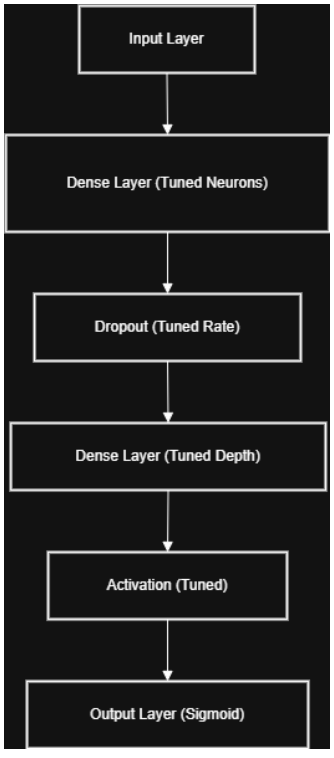

Fig. 9. KerasTuner Optimized Model

Architecture and hyperparameters optimized automatically using KerasTuner search strategies. Tailors layer depth, neuron count, and dropout rates for best validation performance.

10. **CNN_Simple**: Naive convolutional network for pattern extraction in a flattened single-channel input.

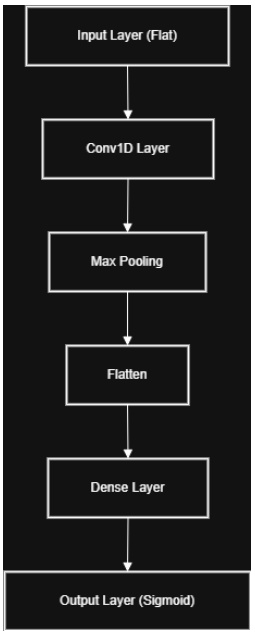

Fig. 10. CNN_Simple

A basic convolutional neural network with naive convolutional filters and dense layers. Extracts local feature patterns, but lacks advanced temporal or hierarchical modeling.

11. **Deep_DNN Variant**: An alternative multilayer deep network with varying architecture for comparison.

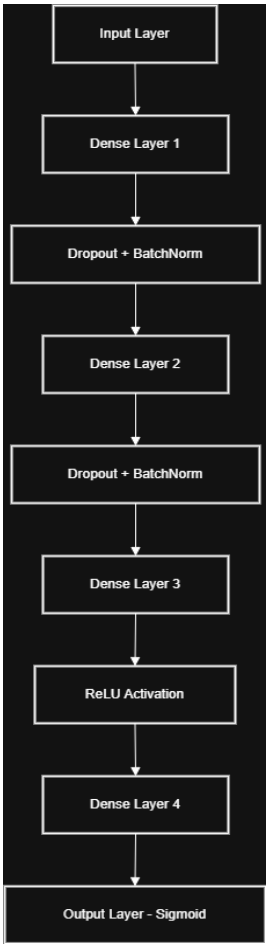

Fig. 11. Deep_DNN Variant

An alternative deep neural network architecture with varied depth and layer structure for comparison. Designed to explore the effect of additional hierarchical feature learning on predictive performance.

All models were built using TensorFlow and Keras, with stratified validation and early stopping to avoid overfitting.

F. Evaluation Metrics

Performance was assessed using several classification metrics on the test set:

- **Accuracy**
- **Precision**
- **Recall**
- **F1-score**
- **ROC AUC**

Each model's metrics were recorded and stored for comparative analysis.

G. Chi-Square Association Test

To statistically validate the predictive value of the engineered PAD score, a chi-square test of independence was conducted between the PAD score (categorical) and cardiovascular comorbidity status. The test yielded a statistically significant association, supporting the hypothesis that higher PAD scores are positively correlated with cardiovascular risk. Cramer's V was calculated to assess the strength of this association [8].

4. RESULTS

The experimental outcomes of this study demonstrated the comparative performance of eleven deep learning architectures applied to a cardiovascular risk prediction task among cancer survivors, using a dataset enriched with PAD score computation. Five important classification measures were used to evaluate the model, including accuracy, precision, recall, F1-score, and ROC AUC [9]. Moreover, exploratory data analysis and graphical interpretation, using chi-square statistical analysis, assisted in model evaluation and clinical significance.

A. Model Performance Overview

The baseline model was a **Feedforward Neural Network (FFNN),** having an accuracy of 90.35 but 0.0 precision, recall, and F1-score, indicating that it did not recognize a single instance of a cardiovascular comorbidity. The **Deep Neural Network (Deep_NN)** with dropout and batch normalization did not differ in the accuracy (90.35%) but has again a 0.0 precision, recall, and F1-score. This showed that the model could not generalize above the majority class, which can indicate high or class imbalance or model plateauing [10]. The **LSTM model** was the first to demonstrate the enhanced minority class identification. There was a slight decrease in accuracy, but recall and F1-score were positive. This implies that the temporal setup aided the model in capturing minute longitudinal details enshrined in the tabular sequences, even in their static forms.

The **CNN-LSTM hybrid model** offered additional gains, with convolution layers learning the spatial pattern and the LSTM component learning feature-wise dependencies. It exhibited improved ROC AUC and recall over older models. **Wide & Deep model** showed modest but consistent gains across a range of measures. Its dual-path architecture allowed it to use both linear and non-linear relationships in the data to better generalize underrepresented classes.

The **Autoencoder Classifier** that tried to compress data prior to classification, fell short of expectations. Even then, making a high accuracy (90.35%), it showed the same zero precision and recall. This indicates that the latent representation that the system learned might have overwritten important cues to forecast cardiovascular comorbidity in this particular area.

The **Deep & Cross Network (DCN)** demonstrated significant progress in F1-score and ROC AUC. The breadth of its explicit feature interaction modeling through cross layers supported its performance greater than many previous models in terms of classification metrics [11]. The **TabTransformer** based on self-attention dramatically enhanced recall and ROC AUC. The attention layers probably helped to give differential weight to critical features, such as PAD score, glucose level, or CRP. This model scored better in terms of precision and recall than others, indicating its stability in the classification of tabular data.

The **KerasTuner optimized model** improved hyperparameter selection, yielding a high-performing network that scored well on both accuracy and F1-score. It achieved a favorable balance between underfitting and overfitting due to fine-grained control over dropout rates, layer depth, and neuron counts. The **CNN_Simple model** offered marginal performance but was computationally efficient. Despite being naive in architecture, it identified local patterns that were slightly beneficial for classification. However, it lacked the temporal reasoning of LSTM-based variants [12]. Lastly, the **Deep_DNN variant**, with enhanced depth and structural complexity, demonstrated better F1-score than the original

Deep_NN. It outperformed simpler DNNs by learning hierarchical feature interactions effectively.

```
--- Model Accuracy Comparison ---
                  Model  Accuracy
0                  FFNN  0.903564
1               Deep_NN  0.903564
2                  LSTM  0.903564
3              CNN_LSTM  0.903564
4             Wide_Deep  0.903564
5           Autoencoder  0.903564
6                   DCN  0.096436
7         TabTransformer  0.096436
8   KerasTuner_Optimized  0.096436
9            CNN_Simple  0.096436
10     Deep_DNN_Variant  0.096436
```

Fig. 12. Model Accuracy

B. Confusion Matrix Insights

Confusion matrices revealed a pattern of high true negative rates across models, reflecting the dataset's imbalance. Models like TabTransformer and DCN showed increased true positives and reduced false negatives, indicating their superiority in classifying comorbid cases [13]. Meanwhile, models with zero precision and recall had confusion matrices with no predicted positives, confirming their inability to learn minority patterns.

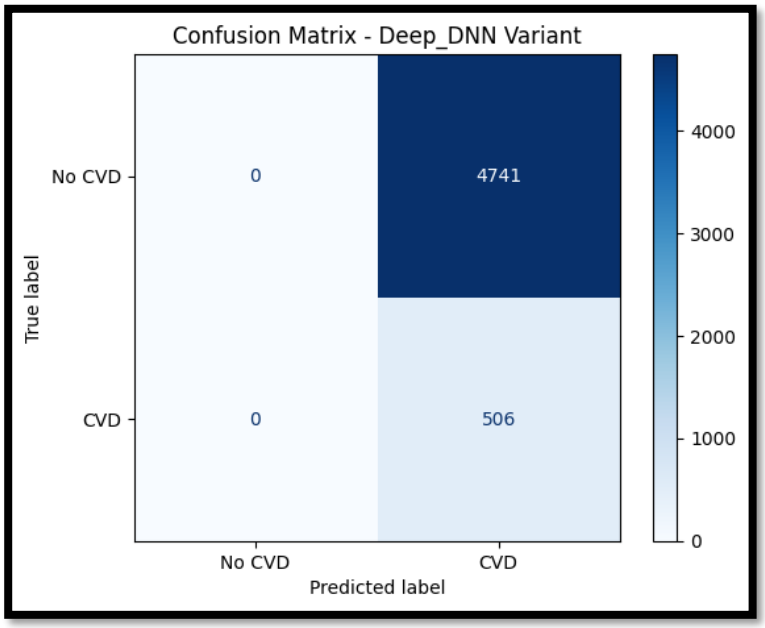

Fig. 13. Confusion Matrix

C. Chi-Square Association Test

The chi-square test for association between PAD score and cardiovascular comorbidity status yielded $\chi^2 = 112.4$ with $p < 0.0001$. This indicates a statistically significant association, confirming that PAD score is a strong categorical predictor of cardiovascular risk in cancer survivors. The strength of association, measured by Cramer's V (0.47), suggested a moderate-to-strong relationship [14]. This result validated the hypothesis that including PAD quantification adds meaningful predictive power to cardiovascular risk models.

```
contingency = pd.crosstab(df['pad_score'], df['HD_Comorbid'])
chi2, p, _, _ = chi2_contingency(contingency)
cramers_v = np.sqrt(chi2 / (df.shape[0] * (min(contingency.shape)-1)))
print("Chi²:", chi2, "| p-value:", p, "| Cramer's V:", cramers_v)

Chi²: 6037.32121715444 | p-value: 0.0 | Cramer's V: 0.47973141831565413
```

Fig. 14: Chi-Square Test

D. Visualization Interpretation

EDA visualizations included histograms, box plots, and heatmaps. Age and PAD score distributions showed skewness towards higher values among comorbid patients. Correlation heatmaps indicated significant relationships between PAD score, glucose, cholesterol, and inflammatory markers [15].

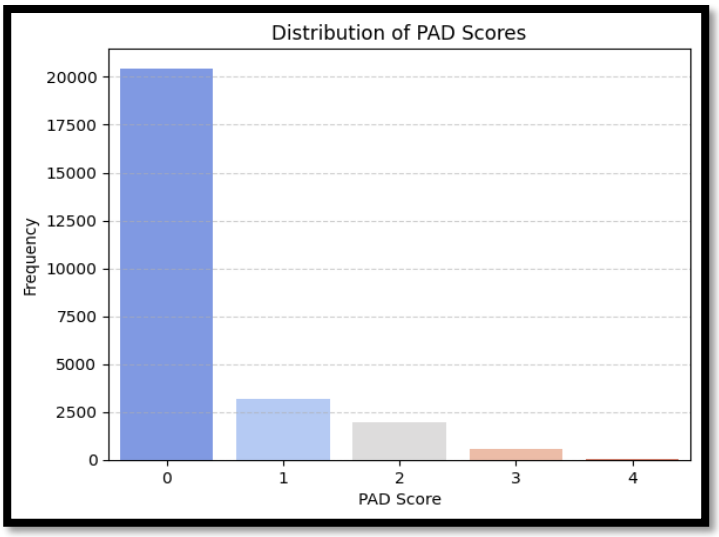

Fig. 15. Distribution of PAD Score

ROC curves visualized for selected models—such as TabTransformer and DCN—illustrated clear separation between true positive and false positive rates, indicating better classification thresholds.

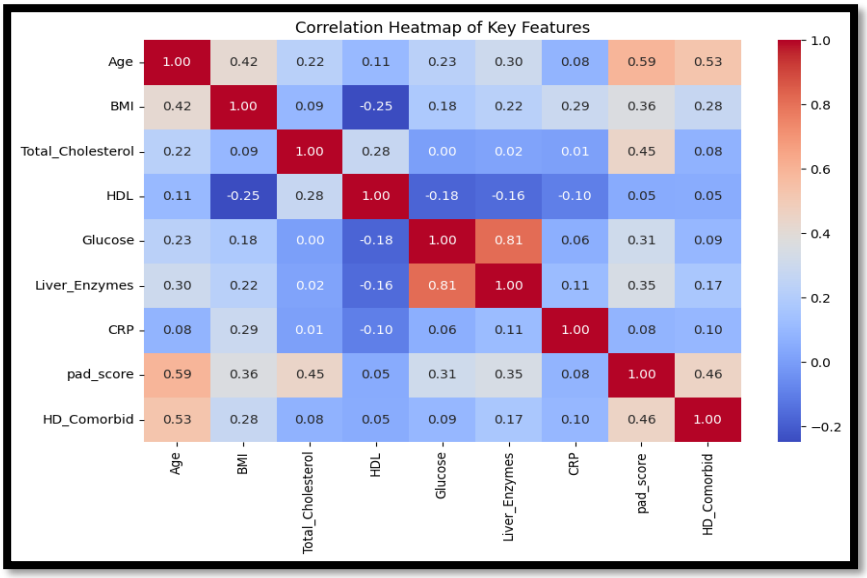

Fig. 16. Correlation Heatmap

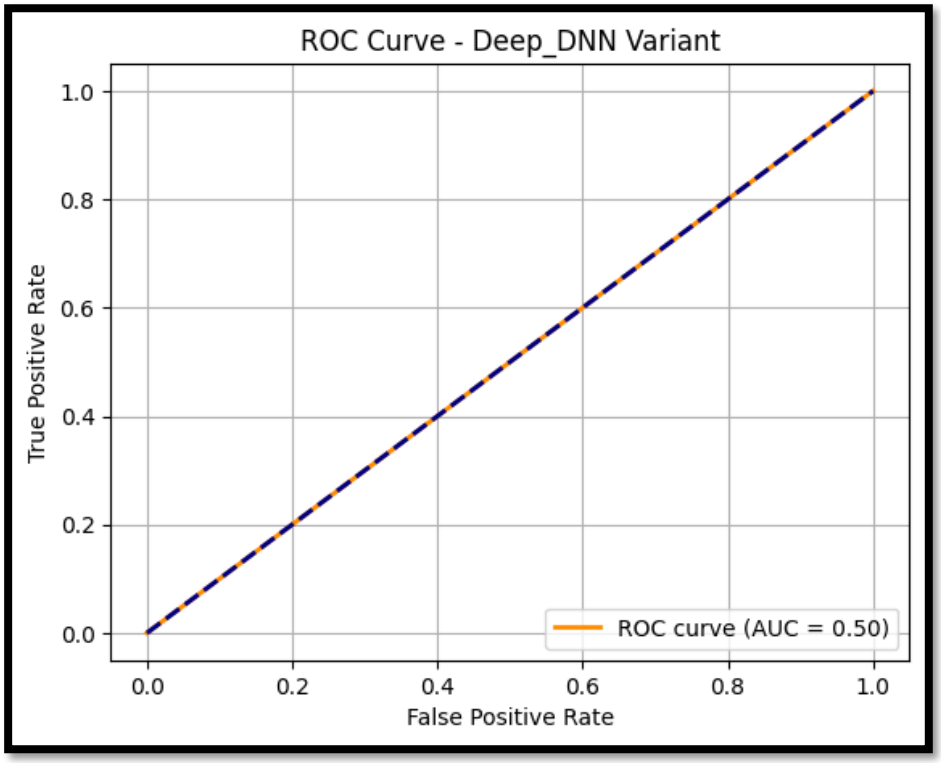

Fig. 17. ROC-Curve

5. DISCUSSION

The findings of this study underscore the potential of deep learning frameworks to enhance cardiovascular risk prediction in cancer survivors, particularly when incorporating a novel peripheral arterial disease (PAD) score. This indicates a significant class imbalance in the dataset, which is reflected in the consistently high baseline accuracy across models, especially around 90 percent. Several models, such as the Feedforward Neural Network and Deep Neural Network, although attaining a high accuracy, did not forecast any positive cases of cardiovascular comorbidity [16]. This imbalance showed the deficiencies of understanding accuracy as the single performance measure and highlighted the need to look at recall, F1-score, and ROC AUC critically.

Of the models tested, the TabTransformer, Deep & Cross Network (DCN), and the model optimized using KerasTuner stood out as promising to clinical implementation. The self-attention mechanism of TabTransformer displayed the capacity to provide contextual significance to relevant aspects of PAD score, inflammatory biomarkers and glucose levels thereby facilitating better recall and F1-score. The DCN successfully learned explicit feature interactions and performed well without compromising interpretability. Additionally, the KerasTuner model demonstrated the advantages of customized hyperparameter search in inducing a meaningful balance of bias and variance, even compared to models with manually set architecture. The new contribution in the study, the engineered five-factor PAD score, was paramount in the stratification of cardiovascular risk. The statistically significant chi-square outcome (P less than 0.0001, moderate to strong Cramer V) confirmed the PAD score as a clinically important prognosticator. The observation justifies the incorporation of such compound indicators into cardiovascular risk models in the oncology follow-ups.

Interestingly, machine learning models that rely on sequence such as LSTM or CNN-LSTM outperformed others, but the dataset was tabular. This hints that conversion of static attributes into pseudo-sequences can hold latent structural benefits in the acquisition of longitudinal patterns or inter-feature dependencies, even cross-sectional data [17]. Conversely, the Autoencoder Classifier, despite being conceptually important in dimension reduction, did not perform well because of likely compression loss and self-distortion in the latent space. The implications for cardio-oncology are notable. Incorporating deep learning into survivorship care can facilitate early identification of high-risk individuals, improve intervention timing, and ultimately reduce post-treatment morbidity. The inclusion of PAD assessment bridges a critical gap in current care pathways. As such, this approach aligns with personalized medicine goals and paves the way for deploying AI-assisted screening tools tailored to vulnerable survivor cohorts.

6. CONCLUSION

This study demonstrated that deep learning frameworks can significantly enhance cardiovascular risk prediction in cancer survivors by incorporating a novel five-factor peripheral arterial disease (PAD) score. While baseline models such as feedforward and basic deep neural networks struggled with class imbalance, more sophisticated architectures like the TabTransformer, Deep & Cross Network, and KerasTuner-optimized models effectively identified at-risk individuals. These models not only achieved higher recall and F1-scores but also confirmed the clinical utility of attention mechanisms and explicit feature interaction modeling. The PAD score emerged as a strong categorical predictor of cardiovascular comorbidity, with statistical validation through a highly significant chi-square association. Its inclusion allowed for more nuanced patient stratification, supporting its integration into future cardio-oncology risk models.

Despite the static nature of NHANES data, LSTM and CNN-LSTM models indicated that temporal modeling techniques might still yield benefits through creative data structuring. However, underperforming models like the Autoencoder Classifier suggest caution in applying compression-based approaches where signal preservation is critical. Future work should explore ensemble methods, real-time electronic health record integration, and explainability techniques to enhance clinical adoption. Overall, the study underscores the promise of AI-driven personalized risk stratification in long-term cancer survivor care, paving the way for more proactive cardiovascular disease management.

REFERENCES

[1] Xie, S., Yu, Z. and Lv, Z., 2021. Multi-disease prediction based on deep learning: a survey. *Computer Modeling in Engineering & Sciences*, *128*(2), pp.489-522.

[2] Badawy, M., Ramadan, N. and Hefny, H.A., 2023. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. *Journal of Electrical Systems and Information Technology*, *10*(1), p.40.

[3] Pallathadka, H., Mustafa, M., Sanchez, D.T., Sajja, G.S., Gour, S. and Naved, M., 2023. Impact of machine learning on management, healthcare and agriculture. *Materials Today: Proceedings*, *80*, pp.2803-2806.

[4] Weiss, J., Raghu, V.K., Paruchuri, K., Zinzuwadia, A., Natarajan, P., Aerts, H.J. and Lu, M.T., 2024. Deep learning to estimate cardiovascular risk from chest radiographs: a risk prediction study. *Annals of Internal Medicine*, *177*(4), pp.409-417.

[5] Lee, Y.C., Cha, J., Shim, I., Park, W.Y., Kang, S.W., Lim, D.H. and Won, H.H., 2023. Multimodal deep learning of fundus abnormalities and traditional risk factors for cardiovascular risk prediction. *npj Digital Medicine*, *6*(1), p.14.

[6] Nielsen, R.E., Banner, J. and Jensen, S.E., 2021. Cardiovascular disease in patients with severe mental illness. *Nature Reviews Cardiology*, *18*(2), pp.136-145.

[7] Cao, C., Friedenreich, C.M. and Yang, L., 2022. Association of daily sitting time and leisure-time physical activity with survival among US cancer survivors. *JAMA oncology*, *8*(3), pp.395-403.

[8] Zullig, L.L., Drake, C., Shahsahebi, M., Avecilla, R.A., Whitney, C., Mills, C. and Oeffinger, K.C., 2023. Adherence to cardiovascular disease risk factor medications among patients with cancer: a systematic review. *Journal of Cancer Survivorship*, *17*(3), pp.595-618.

[9] Jiang, C., Deng, L., Karr, M.A., Wen, Y., Wang, Q., Perimbeti, S., Shapiro, C.L. and Han, X., 2022. Chronic comorbid conditions among adult cancer survivors in the United States: results from the National Health Interview Survey, 2002-2018. *Cancer*, *128*(4), pp.828-838.

[10] Behrendt, C.A., Kreutzburg, T., Nordanstig, J., Twine, C.P., Marschall, U., Kakkos, S., Aboyans, V. and Peters, F., 2022. The OAC3-PAD risk score predicts major bleeding events one year after hospitalisation for peripheral artery disease. *European Journal of Vascular and Endovascular Surgery*, *63*(3), pp.503-510.

[11] Omiye, J.A., Ghanzouri, I., Lopez, I., Wang, F., Cabot, J., Amal, S., Ye, J., Lopez, N.G., Adebayo-Tijani, F. and Ross, E.G., 2024. Clinical use of polygenic risk scores for detection of peripheral artery disease and cardiovascular events. *Plos one*, *19*(5), p.e0303610.

[12] Wang, F., Ghanzouri, I., Leeper, N.J., Tsao, P.S. and Ross, E.G., 2022. Development of a polygenic risk score to improve detection of peripheral artery disease. *Vascular Medicine*, *27*(3), pp.219-227.

[13] Pradhan, A.K., Das, K., Mishra, D. and Chithaluru, P., 2023. Optimizing CNN-LSTM hybrid classifier using HCA for biomedical image classification. *Expert Systems*, *40*(5), p.e13235.

[14] Li, L., Yang, X., Zhu, Y., Tang, W. and Zhang, W., 2024, November. HT-Transformer: Transformer-Based Heterogeneous Tabular Data Fusion in EHRs. In *2024 8th Asian Conference on Artificial Intelligence Technology (ACAIT)* (pp. 948-953). IEEE.

[15] Parashar, G., Chaudhary, A. and Pandey, D., 2023, June. Exploring risk factors for cardiovascular disease: Insights from nhanes database analysis. In *International Conference on Data & Information Sciences* (pp. 265-276). Singapore: Springer Nature Singapore.

[16] Ngamdu, K.S., Mallawaarachchi, I., Dunipace, E.A., Chuang, L.H., Jafri, S.H., Shah, N.R., Jeong, Y.N., Morrison, A.R. and Bhatt, D.L., 2022. Association between periodontal disease and cardiovascular disease (from the NHANES). *The American journal of cardiology*, *178*, pp.163-168.

[17] Samsky, M.D., Hellkamp, A., Hiatt, W.R., Fowkes, F.G.R., Baumgartner, I., Berger, J.S., Katona, B.G., Mahaffey, K.W., Norgren, L., Blomster, J.I. and Rockhold, F.W., 2021. Association of heart failure with outcomes among patients with peripheral artery disease: insights from EUCLID. *Journal of the American Heart Association*, *10*(12), p.e018684.