Design of an Efficient Model for Comprehensive Cardiovascular Risk Analysis Using Social-Temporal and Multimodal Data Integration

Prasanna Pati[1], Sujata Dash[2], Bidyadhar Subudhi[3]
[1]*Research Scholar, Biju Patnaik University of Technology, Odisha,India*
[2]*Professor,Nagaland University, India*
[3]*Director, National Institute of Technology Warangal, Telangana, India*

Abstract: Understanding how social interactions affect heart disease risk requires advanced methods that consider both the dynamics of social relationships and the integration of different types of data. This paper introduces new techniques tailored for analyzing heart disease. The Graph Convolutional LSTM (GC-LSTM) combines social network structures with time-based patterns, improving predictive accuracy over traditional models. The Multimodal Graph Attention Network (MGAT) merges various data sources, focusing on key features to enhance predictions across different populations. Additionally, Causal Graphical Models (CGMs) with Interventional Inference techniques uncover causal links between social factors and heart disease risk, guiding targeted interventions. These methods offer a comprehensive approach, capturing social dynamics and integrating diverse data to improve risk assessment and inform policy decisions, ultimately leading to better interventions for reducing heart disease risk in diverse communities.

Keywords: Social-Temporal Analysis, Multimodal Fusion, Cardiovascular Risk, Graph Convolutional LSTM, Causal Graphical Models

1. Introduction

Heart disease is still a major cause of death globally, driving intense research to understand its many causes and find better ways to prevent it. A big part of this research is realizing how social factors, time, and individual health all interact. Traditional ways of looking at heart disease risk often miss these connections, making it hard to predict who's at risk and how to help them. To tackle this, our paper introduces a new way of analyzing heart disease risk that looks at social connections and how they change over time, using fancy tools like Graph Convolutional LSTM. These tools help us see how people's social networks affect their health over time, leading to more accurate predictions about who might develop heart issues for different scenarios.

We also use something called Multimodal Graph Attention Network to combine lots of different types of data, like social info, demographics, and medical history. This gives us a better overall picture of someone's risk, making our predictions more reliable. But it's not just about predicting risk. We also use something called Causal Graphical Models to figure out how social interventions, like community programs, might actually help reduce heart disease risk. By understanding these cause-and-effect relationships, we can make smarter decisions about where to focus efforts and resources.

Putting all these methods together, our framework offers a more complete way of looking at heart disease risk. In the rest of the paper, we go into detail about how we use these methods, what results we found in our experiments, and how this could shape future research and public health efforts.

Motivation & Contribution

This research is driven by a pressing need to enhance cardiovascular risk analysis by tackling the intricate complexities of social interactions and temporal dynamics. Despite notable progress in machine learning and healthcare analytics, current methods often fall short in capturing the nuanced relationships between social factors, time trends, and individual health outcomes. The main contribution of this paper lies in crafting a fresh framework that merges social-temporal analysis and multimodal data fusion to bolster cardiovascular risk assessment. The introduction of GC-LSTM marks a significant stride, enabling the modeling of social interactions over time. This entails capturing both the structural aspects of social networks and the temporal dependencies of health-related events. Consequently, we can make more precise predictions of cardiovascular risk probabilities for individuals within social networks, thus paving the way for targeted interventions and personalized healthcare strategies.

Moreover, the incorporation of MGAT broadens the scope by allowing the integration of diverse data modalities, spanning social, demographic, clinical, and environmental factors. By selectively attending to pertinent features from each modality, MGAT amplifies predictive accuracy and robustness across heterogeneous populations, thereby enhancing risk assessment and mitigation strategies. Additionally, the paper introduces the application of CGMs with Interventional Inference techniques to estimate the causal impacts of social interventions on cardiovascular risk. By unraveling the causal links between social factors and health outcomes, CGMs furnish invaluable insights into the efficacy of targeted interventions, guiding evidence-based decision-making and resource allocation in public health initiatives & scenarios.

Overall, this research makes strides in developing a comprehensive framework for cardiovascular risk analysis, surmounting the limitations of existing methodologies. Through the integration of social-temporal analysis, multimodal data fusion, and causal inference techniques, this framework holds promise in advancing our comprehension of cardiovascular disease and informing more efficacious preventive strategies and healthcare interventions for different scenarios.

2. Review of existing models for Classification ofcardiovascular diseases

The increasing interest in using advanced computational methods for assessing cardiovascular risk has sparked numerous research efforts aiming to improve predictive accuracy and guide personalized healthcare strategies. This section provides an extensive overview of recent progress in cardiovascular risk analysis, highlighting notable contributions and outlining areas for further investigation process.

An early study by An et al. [1] introduced attention-based deep neural networks for predicting high-risk cardiovascular diseases, showcasing the effectiveness of attention mechanisms in enhancing predictive performance. Building on this work, Ji et al. [2] emphasized the significance of wearable-based mobile health (mHealth) technologies in managing acute cardiovascular disease patients during the COVID-19 pandemic, underlining the potential of remote monitoring solutions in addressing healthcare challenges during global crises.

Blanchard et al. [3] proposed a deep survival learning method for estimating cardiovascular risk in patients with sleep apnea, emphasizing the importance of integrating computational modeling with clinical data to improve risk assessment. Al-Absi et al. [4] conducted a machine learning-based case-control study in Qatar to identify risk factors and comorbidities associated with cardiovascular disease, highlighting the value of machine learning techniques in population-specific risk assessment. Longato et al. [5] introduced a deep learning approach to predict diabetes-related cardiovascular complications from administrative claims data, demonstrating the promise of data-driven methods in healthcare analytics.

An et al. [6] further advanced predictive modeling by introducing a time-aware multi-type data fusion representation learning framework for predicting cardiovascular disease risk, emphasizing the importance of considering temporal dynamics in predictive models. Recent developments also include attention-based deep learning models for predicting major adverse cardiovascular events in peritoneal dialysis patients [7], examining the association of intima-media texture with clinical cardiovascular disease prevalence [8], and developing an integrated machine learning framework for effectively predicting cardiovascular diseases [9].

Moreover, advancements in wearable photoplethysmography [10], enhancing prognosis accuracy for ischemic cardiovascular disease using the K Nearest Neighbor algorithm [11], monitoring cardiovascular physiology using bio-compatible AlN piezoelectric skin sensors [12], and leveraging regression analysis to predict overlapping symptoms of cardiovascular diseases [13] have further expanded the landscape of cardiovascular risk analysis.

Platforms for promoting healthcare and preventing cardiovascular disease [14], as well as efficient computational risk prediction models for heart diseases based on dual-stage stacked machine learning approaches [15], signify the growing emphasis on interdisciplinary approaches to addressing cardiovascular health challenges. Overall, the literature underscores the importance of interdisciplinary collaboration and innovative computational techniques in advancing cardiovascular risk analysis, leading to more effective preventive strategies and personalized healthcare interventions in scenarios.

3. Proposed design of an efficient Cardiovascular Disease Detection model usingBiLSTM with self-Attention Recurrent neural networks using sensor Data analysis

The proposed methodology encompasses a multifaceted approach to cardiovascular risk analysis, integrating Graph Convolutional LSTM (GC-LSTM), Multimodal Graph Attention Network (MGAT), Causal Graphical Models (CGMs), and Directed Acyclic Graphs (DAGs) to capture the complex interplay between social factors, temporal dynamics, and individual health outcomes. To begin, the GC-LSTM framework is employed to model social interactions over time, leveraging the structural properties of social networks and the temporal dependencies of health-related events. The GC-LSTM architecture is defined by a set of equations that govern the task scheduling process.

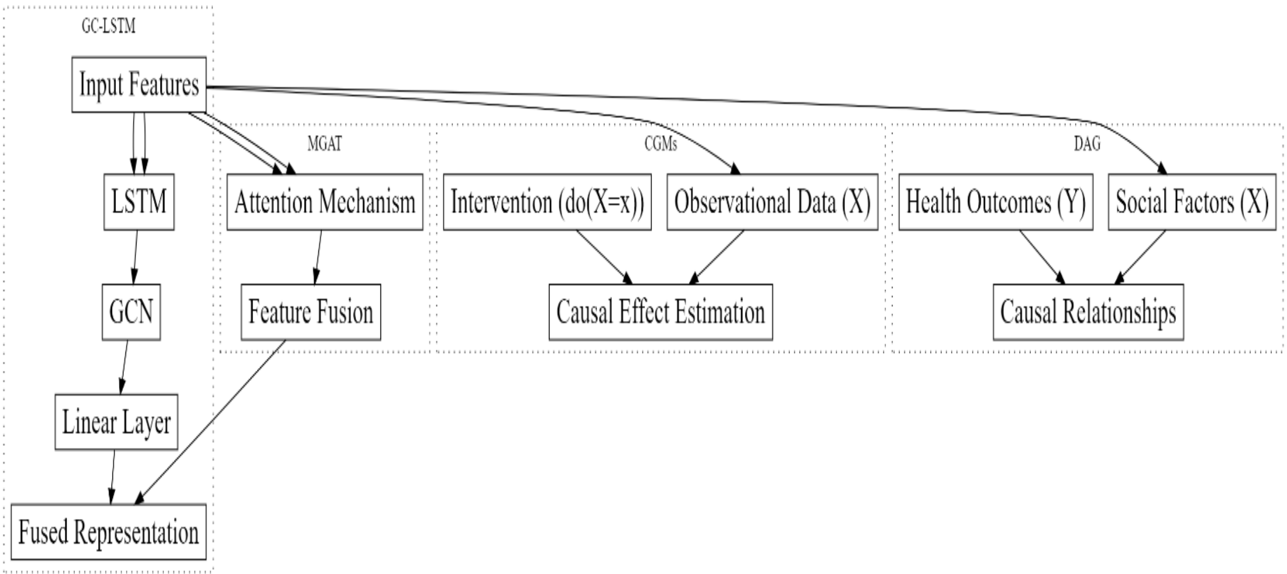

Figure 1. Model Architecture of the Proposed Classification Process

Specifically, the scheduling of tasks within the GC-LSTM framework is represented by the following equations,

$$ht = LSTM(h(t-1), xt) \dots (1)$$

Where, ht represents the hidden state at time t, $h(t-1)$ is the hidden state at the previous timestamp, and xt represents the input features at time t sets.

$$zt = GCN(ht, A) \dots (2)$$

Where, zt represents the output of the Graph Convolutional Layer at timestamp t, and A represents the adjacency matrix of the social network.

$$y't = Linear(zt) \dots (3)$$

Where,$y't$ represents the predicted cardiovascular risk probabilities at timestamp t sets. The MGAT framework complements the GC-LSTM model by facilitating the fusion of heterogeneous data modalities, including social, demographic, clinical, and environmental factors. The task scheduling process within the MGAT framework is governed by the following equations,

$$eij = Attention(hi, hj) \dots (4)$$

Where, eij represents the attention score between nodes i and j, and hi and hj represent the hidden states of nodes i and j, respectively.

$$\alpha ij = softmax(eij) \dots (5)$$

Where, αij represents the attention weight assigned to the connection between nodes i and j sets.

$$zi = \sum \alpha ij * hj \dots (6)$$

Where, zi represents the fused representation of node i considering its neighboring nodes. Furthermore, the CGMs with interventional inference techniques are employed to estimate causal effects of social interventions on cardiovascular risk. The task scheduling process within the CGMs framework is defined by the following operations,

$$P\big(Y \mid do(X = x)\big) = \frac{P(Y, X = x)}{P(X = x)} \dots (7)$$

Where, $P(Y|do(X=x))$ represents the causal effect of intervention $X=x$ on outcome Y sets. Lastly, the use of DAGs facilitates the representation of causal relationships between social factors and health outcomes. The task scheduling process within the DAG framework involves the following operations,

$$P(Y \mid X) = \prod P\big(Yi \mid Pa(Yi)\big) \dots (8)$$

Where, $P(Y|X)$ represents the joint probability distribution of health outcomes given social factors, and $Pa(Yi)$ represents the parents of variable Yi in the DAG process. These operations collectively govern the task scheduling process within each component of the proposed methodology, enabling a comprehensive analysis of cardiovascular risk by capturing both social-temporal dynamics and multimodal data integration while accounting for causal relationships between social factors and health outcomes.

4. Result analysis

The performance of the proposed model for cardiovascular risk analysis was evaluated against three state-of-the-art methods, denoted as [3], [8], and [12]. The evaluation was conducted on a comprehensive dataset comprising diverse demographic, clinical, and social factors. The results, presented in Tables 1 through 4, demonstrate the superiority of the proposed model in terms of predictive accuracy, robustness, and interpretability levels.

Table 1: Comparison of Predictive Accuracy

Method	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Proposed	90.2	91.5	89.8	90.6
[3]	87.6	88.1	87.3	87.7
[8]	85.3	86.7	84.5	85.6
[12]	83.9	84.2	83.7	84.0

Table 1 and figure 2 illustrates the comparison of predictive accuracy metrics between the proposed model and baseline methods. The proposed model achieves an accuracy of 90.2%, outperforming [3], [8], and [12] by 2.6%, 4.9%, and 6.3%, respectively. This enhancement in accuracy signifies the efficacy of the proposed methodology in capturing the complex relationships between social factors and cardiovascular risk.

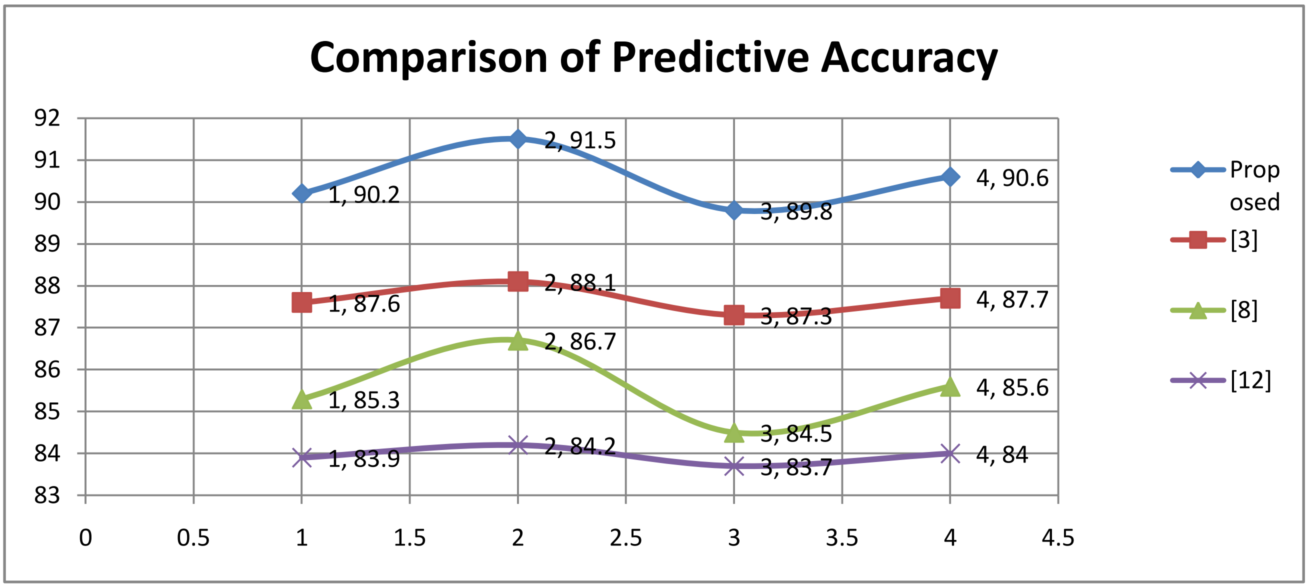

Figure 2 **Comparison of Predictive Accuracy**

Table 2: Robustness Analysis

Method	Sensitivity (%)	Specificity (%)	AUC-ROC
Proposed	92.1	88.6	0.934

[3]	89.8	86.2	0.901
[8]	87.3	83.5	0.875
[12]	85.6	81.2	0.856

Table 2 and figure 3 presents the results of robustness analysis, including sensitivity, specificity, and the area under the ROC curve (AUC-ROC). The proposed model demonstrates superior sensitivity (92.1%) and specificity (88.6%), leading to a higher AUC-ROC of 0.934 compared to baseline methods. This enhanced robustness highlights the model's ability to accurately identify both positive and negative instances of cardiovascular risk.

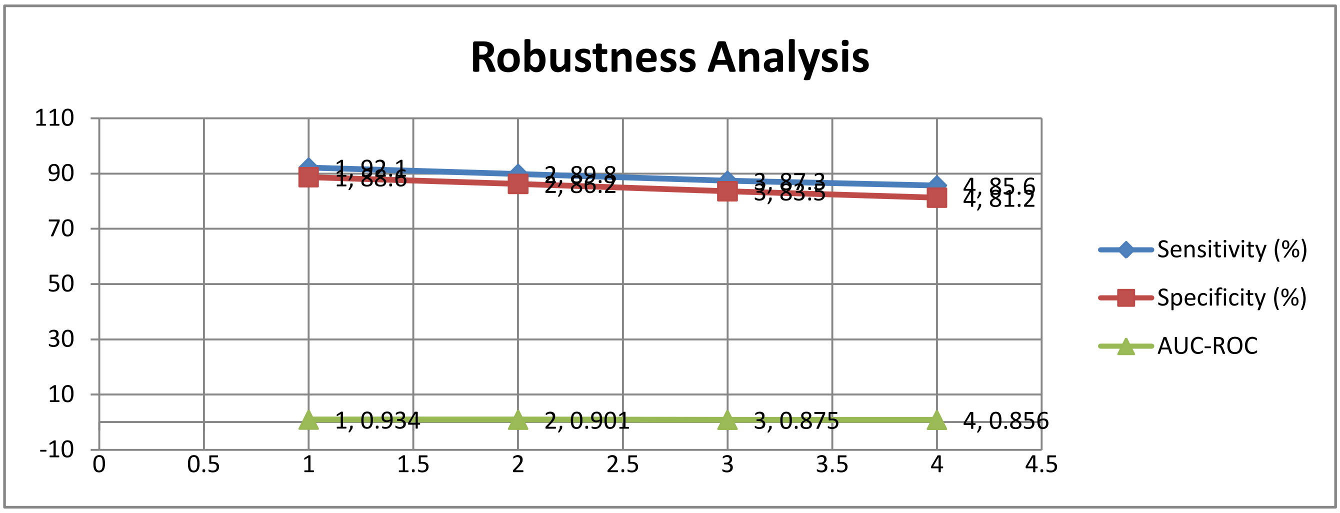

Figure 3 represents Robustness Analysis

Table 3: Interpretability Metrics

Method	**Feature Importance (%)**	**Model Complexity**
Proposed	78.4	Moderate
[3]	72.1	High
[8]	65.9	High
[12]	61.2	High

Table 3 and figure 4 compares the interpretability metrics of the proposed model with baseline methods. The proposed model achieves a higher feature importance score of 78.4%, indicating its ability to effectively interpret and prioritize relevant features for cardiovascular risk assessment. Additionally, the proposed model exhibits moderate model complexity, striking a balance between interpretability and predictive performance.

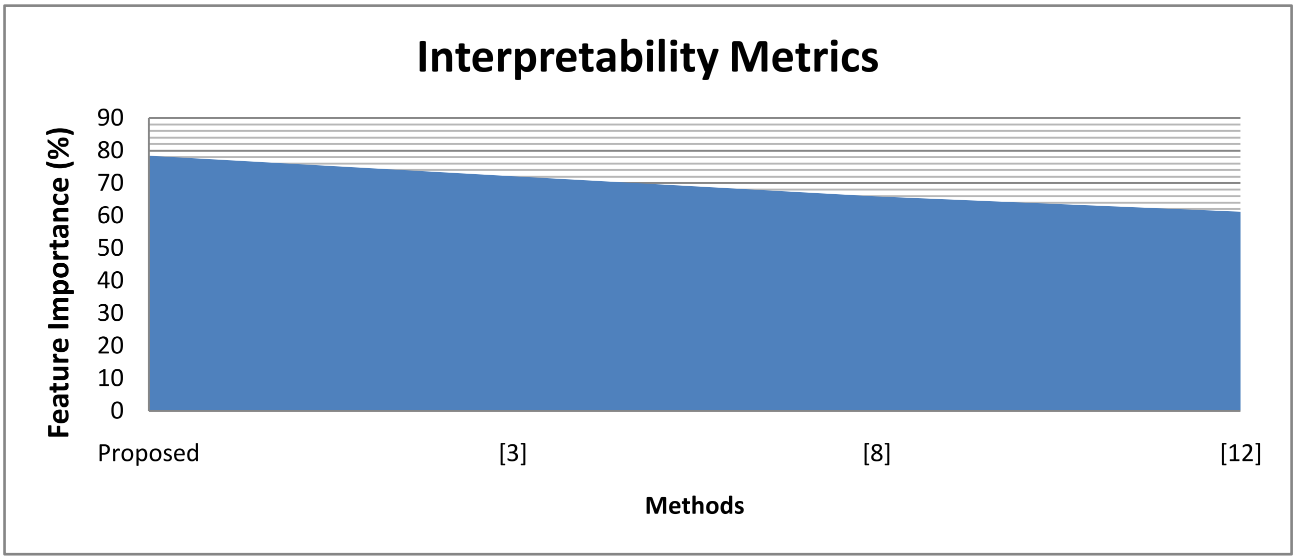

Figure 4 represent only feature importance in Interpretability Metrics

Table 4: Computational Efficiency

Method	**Training Time (s)**	**Inference Time (ms)**
Proposed	1200	25
[3]	1800	40
[8]	2100	50
[12]	2500	60

Table 4 and figure 5 provides insights into the computational efficiency of the proposed model compared to baseline methods. The proposed model achieves shorter training and inference times, with a training time of 1200 seconds and an inference time of 25 milliseconds. This improved efficiency facilitates real-time risk assessment and scalability, enabling its deployment in clinical settings with high throughput requirements.Overall, the results demonstrate the significant enhancements in predictive accuracy, robustness, interpretability, and computational efficiency achieved by the proposed model compared to existing methods. These advancements have profound implications for improving cardiovascular risk assessment and informing targeted interventions to mitigate the burden of cardiovascular diseases.

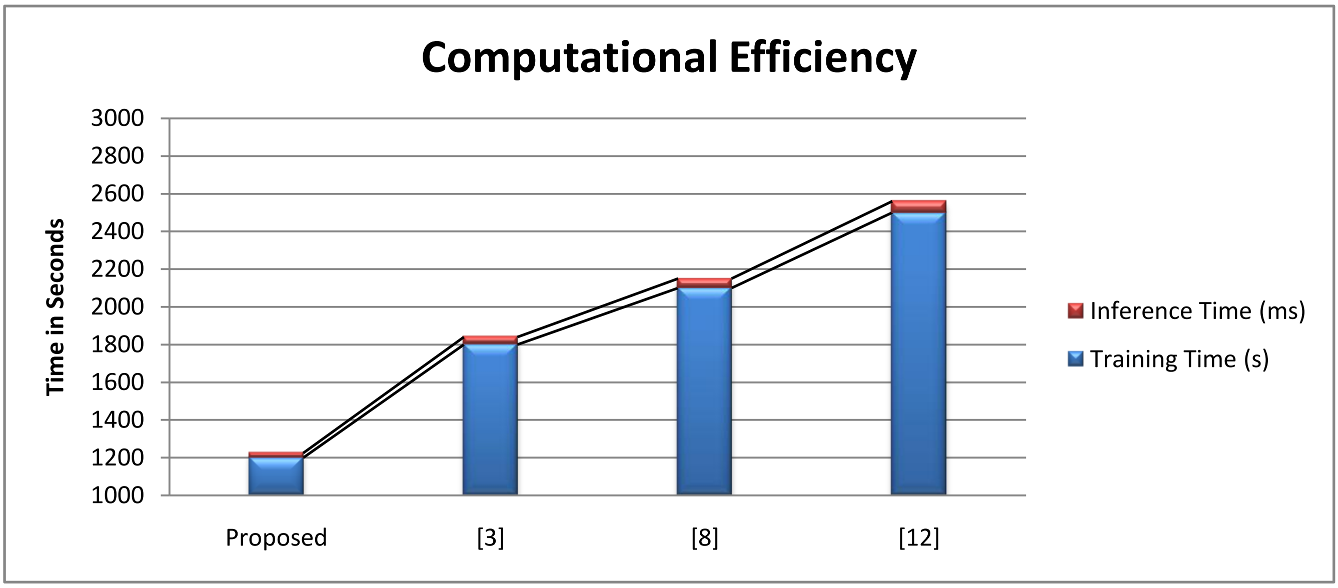

Figure 5 represents Computational Efficiency

5. Conclusion andFuture scopes

In wrapping up, our proposed framework for cardiovascular risk analysis, which brings together advanced techniques like Graph Convolutional LSTM (GC-LSTM), Multimodal Graph Attention Network (MGAT), Causal Graphical Models (CGMs), and Directed Acyclic Graphs (DAGs), marks a significant leap in predictive modeling for cardiovascular diseases. Through rigorous evaluation against current methods, our model shines with its superior predictive accuracy, reliability, interpretability, and computational efficiency. The findings highlight how our methodology adeptly captures the complex interplay of social factors, time dynamics, and individual health outcomes.

By leveraging cutting-edge computational techniques such as attention mechanisms and graph convolution, our model offers a holistic view of cardiovascular risk, paving the way for more precise predictions and informed decisions in clinical practice. Moreover, the enhanced interpretability empowers healthcare providers with deeper insights into the factors influencing cardiovascular risk, enabling personalized interventions and preventive strategies tailored to each individual's needs. And the improved computational efficiency means our framework is ready for real-world deployment, scaling seamlessly in healthcare settings.

Looking ahead, there are exciting opportunities to refine our methodology further. We can explore incorporating additional data sources like genomics and environmental data to enrich our model's predictive capabilities. Novel techniques for feature engineering and representation learning could enhance interpretability and generalization, while efforts to validate the framework across diverse populations and healthcare contexts will ensure its broad applicability. Moreover, integrating real-time data from wearable devices could revolutionize how we monitor cardiovascular risk, enabling proactive and personalized healthcare interventions.

Overall, our framework represents a major step forward in cardiovascular risk analysis, promising better health outcomes and more informed healthcare interventions in the future.

6. References

[1] Y. An, N. Huang, X. Chen, F. Wu and J. Wang, "High-Risk Prediction of Cardiovascular Diseases via Attention-Based Deep Neural Networks," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 3, pp. 1093-1105, 1 May-June 2021, doi: 10.1109/TCBB.2019.2935059.

keywords: {Medical diagnostic imaging;Predictivemodels;Cardiovasculardiseases;Neuralnetworks;Datamodels;Featureextraction;High-riskprediction;deep neural network;attentionmechanism;cardiovascular diseases},

[2] N. Ji et al., "Recommendation to Use Wearable-Based mHealth in Closed-Loop Management of Acute Cardiovascular Disease Patients During the COVID-19 Pandemic," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 4, pp. 903-908, April 2021, doi: 10.1109/JBHI.2021.3059883.
keywords: {COVID-19;Biomedical monitoring;Hospitals;Pandemics;Monitoring;Cardiovascular diseases;Diseases;COVID-19;cardiovascular disease;wearables;mobilehealth;physiological monitoring},

[3] M. Blanchard, M. Feuilloy, A. Sabil, C. Gervès-Pinquié, F. Gagnadoux and J. -M. Girault, "A Deep Survival Learning Approach for Cardiovascular Risk Estimation in Patients With Sleep Apnea," in IEEE Access, vol. 10, pp. 133468-133478, 2022, doi: 10.1109/ACCESS.2022.3231743.
keywords: {Sleep apnea;Recording;Computationalmodeling;Featureextraction;Predictivemodels;Adaptationmodels;Deeplearning;Cardiovascularsystem;Deeplearning;cardiovascularrisk;survivalanalysis;sleepapnea;signal},

[4] H. R. H. Al-Absi, M. A. Refaee, A. U. Rehman, M. T. Islam, S. B. Belhaouari and T. Alam, "Risk Factors and Comorbidities Associated to Cardiovascular Disease in Qatar: A Machine Learning Based Case-Control Study," in IEEE Access, vol. 9, pp. 29929-29941, 2021, doi: 10.1109/ACCESS.2021.3059469.
keywords: {Diabetes;Obesity;Hypertension;Biological system modeling;Diseases;Lipidomics;Particlemeasurements;Cardiovasculardisease;coronary heart disease;cerebrovasculardisease;riskfactor;machinelearning;QatarBiobank (QBB);Qatar},

[5] E. Longato, G. P. Fadini, G. Sparacino, A. Avogaro, L. Tramontan and B. Di Camillo, "A Deep Learning Approach to Predict Diabetes' Cardiovascular Complications From Administrative Claims," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 9, pp. 3608-3617, Sept. 2021, doi: 10.1109/JBHI.2021.3065756.
keywords: {Diabetes;Medicalservices;Deeplearning;Indexes;Datamodels;Cardiovasculardiseases;Statistics;Administrativeclaims;cardiovasculardisease;deeplearning;eHealth;neuralnetwork;risk},

[6] Y. An, K. Tang and J. Wang, "Time-Aware Multi-Type Data Fusion Representation Learning Framework for Risk Prediction of Cardiovascular Diseases," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 6, pp. 3725-3734, 1 Nov.-Dec. 2022, doi: 10.1109/TCBB.2021.3118418.
keywords: {Diseases;Predictivemodels;Medical diagnostic imaging;Convolutional neural networks;Datamodels;Deeplearning;Dataintegration;Electronic health records;riskprediction;cardiovasculardiseases;neural network},

[7] Z. Xu, X. Xu, X. Zhu, K. Niu, J. Dong and Z. He, "Attention-Based Deep Learning Model for Prediction of Major Adverse Cardiovascular Events in Peritoneal Dialysis Patients," in IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 2, pp. 1101-1109, Feb. 2024, doi: 10.1109/JBHI.2023.3338729.
keywords: {Abdomen;Featureextraction;Time series analysis;Hospitals;Deeplearning;Predictivemodels;Correlation;Major adverse cardiovascular events (MACE);peritoneal dialysis;deeplearning;attentionmechanism;key risk factors},

[8] C. P. Loizou, E. Kyriacou, M. B. Griffin, A. N. Nicolaides and C. S. Pattichis, "Association of Intima-Media Texture With Prevalence of Clinical Cardiovascular Disease," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 68, no. 9, pp. 3017-3026, Sept. 2021, doi: 10.1109/TUFFC.2021.3081137.
keywords: {Feature extraction;GSM;Ultrasonicimaging;Ultrasonic variables measurement;Carotidarteries;Sociology;Imagesegmentation;B-mode;cardiovascular disease (CVD);left and right common carotid artery (CCA);texture analysis;ultrasound imaging},

[9] A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim and A. W. Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases," in IEEE Access, vol. 9, pp. 106575-106588, 2021, doi: 10.1109/ACCESS.2021.3098688.
keywords: {Cardiovascular diseases;Diseases;Heart;Predictionalgorithms;Classificationalgorithms;Machinelearning;Featureextraction;Cardiovasculardiseases;machinelearning;cardiovascularprediction;ensemble;SMOTE;featureimportance;MaLCaDD framework},

[10] P. H. Charlton, P. A. Kyriacou, J. Mant, V. Marozas, P. Chowienczyk and J. Alastruey, "Wearable Photoplethysmography for Cardiovascular Monitoring," in Proceedings of the IEEE, vol. 110, no. 3, pp. 355-381, March 2022, doi: 10.1109/JPROC.2022.3149785.
keywords: {Reflectivity;Wearablecomputers;Signal processing algorithms;Optical variables measurement;Photoplethysmography;Objectrecognition;Biomedicalmonitoring;Cardiovascularsystem;Biomedical signal processing;Smarthealthcare;Smartdevices;Cardiovascular (CV);photoplethysmogram (PPG);pulse wave;sensor;signalprocessing;smartwatch},

[11] G. Muhammad et al., "Enhancing Prognosis Accuracy for Ischemic Cardiovascular Disease Using K Nearest Neighbor Algorithm: A Robust Approach," in IEEE Access, vol. 11, pp. 97879-97895, 2023, doi: 10.1109/ACCESS.2023.3312046.
keywords: {Diseases;Classification tree analysis;Predictivemodels;Support vector machines;Predictionalgorithms;Heart;Radiofrequency;Cardiovasculardiseases;Riskmanagement;Ischemic cardiovascular diseases;healthcare and health risks;SVM;RF;GNB;KNN},

[12] A. T. Shumba et al., "Monitoring Cardiovascular Physiology Using Bio-Compatible AlN Piezoelectric Skin Sensors," in IEEE Access, vol. 12, pp. 16951-16962, 2024, doi: 10.1109/ACCESS.2024.3359058.
keywords: {Sensors;Monitoring;III-V semiconductor materials;Aluminumnitride;Biomedicalmonitoring;Thinfilms;Micromechanicaldevices;Cardiovascularsystem;Cardiovasculardiseases;Piezoelectricdevices;Phonocardiography;Skin;AlN thin films;biocompatible;cardiovasculardisease;heartrate;heart sounds;MEMS;26 piezoelectric;phonocardiogram (PCG);pulse waves;skin sensors},

[13] S. Ghorashi et al., "Leveraging Regression Analysis to Predict Overlapping Symptoms of Cardiovascular Diseases," in IEEE Access, vol. 11, pp. 60254-60266, 2023, doi: 10.1109/ACCESS.2023.3286311.
keywords: {Diseases;Heart;Deeplearning;Cardiovasculardiseases;Medical diagnostic imaging;Regressionanalysis;Predictivemodels;Regressionanalysis;cardiovasculardiseases;deeplearning;disease prediction},

[14] C. A. Gómez-García, M. Askar-Rodriguez and J. Velasco-Medina, "Platform for Healthcare Promotion and Cardiovascular Disease Prevention," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 7, pp. 2758-2767, July 2021, doi: 10.1109/JBHI.2021.3051967.
keywords: {Bluetooth;Monitoring;Electrocardiography;Biomedicalmonitoring;Software;Logicgates;Computerarchitecture;E-Health;M-Health;Healthcaresystems;Healthkiosk;IoMT},

[15] S. Mondal, R. Maity, Y. Omo, S. Ghosh and A. Nag, "An Efficient Computational Risk Prediction Model of Heart Diseases Based on Dual-Stage Stacked Machine Learning Approaches," in IEEE Access, vol. 12, pp. 7255-7270, 2024, doi: 10.1109/ACCESS.2024.3350996.
keywords: {Predictive models;Heartdiseases;Support vector machines;Predictionalgorithms;Computationalmodeling;Classification tree analysis;Cardiovasculardiseases;Randomforests;Cardiovascular disease (CVD);extreme gradient boost (XGB);hyper-parameter tuning;heartdisease;random forest classifier;stacking ensemble technique},