*Fuzzy Systems and Data Mining IX*
*A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)*

*doi:10.3233/FAIA231105*

# Interpretable Deep Learning Model for Identifying the Immediate Risk of Myocardial Infarction Complications

Mengxiao WANG[a,b], Yanrui JIN[a,b,1], Zhiyuan LI[a,b], Xiaoyang WEI[a,b], Chengliang LIU[a,b,2] and Xiaoxue YANG[c]
[a] *State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[b] *MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai 200240, China*
[c] *Shanghai University of Finance and Economics, Shanghai200433, China*

**Abstract.** Electronic health records (EHR) encompass extensive personal information, diagnostic records, and medical history, enabling the prediction of disease occurrence and mortality risk. The objective of this study is to predict myocardial infarction complications and assess the risk of death by comparing the performance of various deep learning models and traditional machine learning approaches. The findings demonstrate similar performance between two kinds of models in predicting complication. The DeepFM model is commonly employed for Click-through rate (CTR) prediction. To the best of our knowledge, this is the first application of the DeepFM model to the EHR domain, and we have demonstrated its exceptional predictive performance, achieving the accuracy of 93.95%. Moreover, we further classify samples into low, intermediate, and high-risk categories with high confidence. To comprehend these results, we conduct an interpretability analysis of the models' predictions employing SHAP values. This analysis involves ranking the significant features, and summarizing ECG-related features, which hold clinical decision-making revelance for clinicians.

**Keywords.** DeepFM, multi-label classification, lethal risk assessment, SHAP value

## 1. Introduction

Coronary artery disease (CAD), also referred to coronary heart disease, is currently the leading cause of death among adults worldwide, and this trend is expected to continue over the next decade. However, despite many patients who die from coronary heart disease being previously diagnosed and treated, over 50% of sudden cardiac death cases occur without any clear history of coronary artery disease [1]. Current risk prediction models employed in conventional medicine for coronary heart disease, such as the Framingham Risk Score [1] and the Charlson Comorbidity Index [2] for cancer.

[1] Corresponding Author, Yanrui JIN, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; E-mail: jinyanrui@sjtu.edu.cn.
[2] Corresponding Author, Chengliang LIU, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; E-mail: chlliu@sjtu.edu.cn.

These indicators primarily address the long-term population-level risk of coronary heart disease. However, when it is not possible to measure biochemical markers such as troponin due to delayed presentation, data collected by wearable detection devices, including demographic information, electronic health records (EHR), and ECG patterns/conclusions, can be used to initially assess the patient's risk level for coronary heart disease.

In this work, we implement two main tasks. Firstly, we construct a multi-label classification model, and determine the ranking of influential features in predicting myocardial complications. Secondly, considering the unique characteristics of tabular data, which consists of highly sparse and heterogeneous (categorical-continuous-mixed) features, small sample sizes, and extreme values [3], we explore a new deep learning method DeepFM to predict the risk of death for patients and compare it with conventional machine learning methods. To the best of our knowledge, the DeepFM model is commonly used for Click-through rate (CTR) prediction, but few researchers have applied it in the disease risk prediction. This study facilitates the assessment of disease risks in cardiac patients and provides insights into the trajectory of disease prognosis.

## 2. Methodology

### *2.1. Multi-label Strategy*

The objective of multi-label classification is to train a model on the training set $D = \{(X_i, Y_i) \mid 1 \le i \le m\}$ to represent the function mapping $h : \chi \to 2^{\gamma}$ [4]. The multi-label classification problem is transformed into an optimization problem, where we seek to determine the weight values that minimize the average loss function for each label as expressed in Equation (1).

$$\arg\min_{\beta \in \mathbb{R}^n} \frac{1}{qm} \sum_{j=1}^{q} \sum_{i=1}^{m} L(y_{ik}, \hat{y}_{ik}) \tag{1}$$

### *2.2. DeepFM*

Deep learning architectures are specifically designed to incorporate inductive biases that align with the invariances and spatial dependencies present in the data. However, identifying the corresponding invariances in tabular data is challenging, making it difficult to find state-of-the-art deep learning models for prediction.

Linear models assume independence among individual features, disregarding the interrelationships between them. However, in practice, correlations among a large number of features are common. Moreover, medical electronic health records often contain high-dimensional sparse matrices, and directly modeling these matrices can result in heavy computational burdens and slow updates of feature weights. Factorization Machines (FM) have the advantage of addressing both issues, as represented by Equation (2).

$$y_{FM} = \langle w, x \rangle + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \langle V_i, V_j \rangle x_i \cdot x_j \tag{2}$$

DeepFM model combines the FM model with a deep neural network[5]. As shown in **Figure 1**, the Wide & deep architecture of DeepFM includes an embedding stage and a feature interaction stage, allowing for simultaneous extraction of low and high dimensional features.

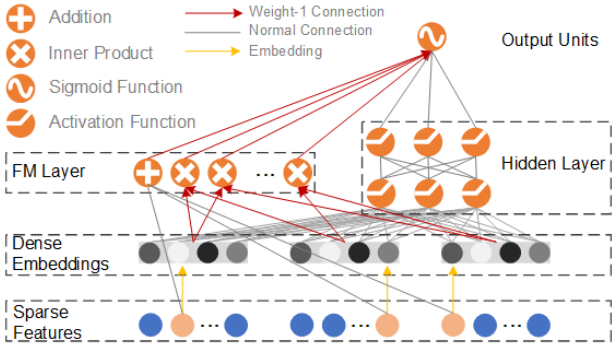


**Figure 1.** Network architecture of DeepFM.

## 3. Experiment Setups

### *3.1. Dataset*

The Myocardial infarction complications Database (MIC)[6] was collected from Krasnoyarsk Interdistrict Clinical Hospital №20 named after I. S. Berzon (Russia) in 1992-1995. The dataset comprises records of 1700 patients, including demographic information, medical history records, electrocardiogram measurements et al. The input features are heterogeneous, which contain 99 categorical features and 12 numeric features. The target labels consist of 11 types of complications and 1 lethal (death) outcome.

### *3.2. Preprocessing*

Initially, a primary data cleaning strategy is implemented, which consists of two steps: removing features and samples with a high missing ratio. The present study employed the Predictive Mean Matching (PMM) imputation method to address the issue of missing values. Since tabular data often consists of both binary and numeric values, it is crucial to standardize the numeric features with Z-score normalization method.

### *3.3. Implementation Details*

In the training process, the Matthews Correlation Coefficient (MCC) is utilized to choose the best threshold. Subsequently, a cross-validated grid search strategy is employed to select the optimal parameters for the machine learning model. The AUC ROC is chosen as the indictor to evaluate the performance of the cross-validated model. We use *scikit-learn*, *pytorch* and *deepctr_torch* to implement the models. All neural network architectures are trained by Adagrad optimization with binary cross-entropy loss.

## 4. Results

Applying different machine learning method to the tabular data, the classification results can be evaluated through different metrics. Results of multi-label classification are revealed in Table 1, which traditional machine learning methods perform better in F1 macro and Jaccard macro. However, for Hamming loss and Accuracy, the NN model stands out.

**Table 1.** Evaluation results of multi-label classification

| Model | Hamming loss | F1 macro | Jaccard macro | Accuracy |
|---|---|---|---|---|
| BR-MultiNB | 0.1386 | **0.2583** | **0.1548** | 0.2939 |
| CC-MultiNB | 0.1427 | 0.2483 | 0.1500 | 0.2909 |
| LP-MultiNB | 0.0967 | 0.1778 | 0.1092 | 0.3667 |
| MLKNN | 0.0914 | 0.0649 | 0.0379 | 0.3697 |
| MLARAM | 0.1265 | 0.0717 | 0.0446 | 0.1364 |
| BRKNNa | 0.1258 | 0.1082 | 0.0625 | 0.2394 |
| BRKNNb | 0.4687 | 0.1160 | 0.0644 | 0.0303 |
| MLSVM | 0.1682 | 0.2061 | 0.1200 | 0.1576 |
| NN(MLP) | **0.0788** | 0.0732 | 0.1024 | **0.4061** |

The prediction results of whether patients have lethal risk of myocardial infarction is depicted in Figure 2. Each value represents the test score of the best model (on the validation set) obtained after a specific number of iterations in a random search. These values are extracted over 5 shuffles of the random search order. In comparison, DeepFM performs excellent with accuracy of 0.9395.

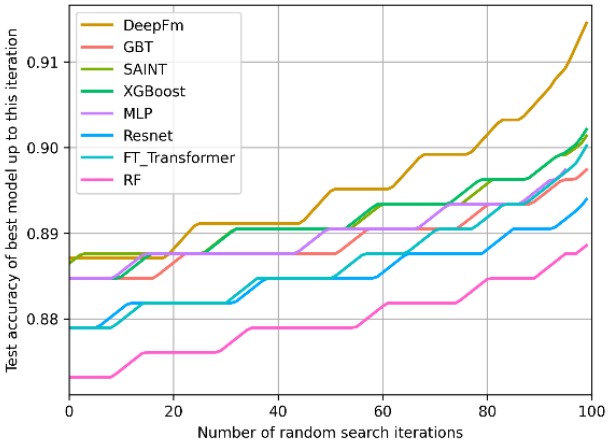


**Figure 2.** Test accuracy of predicting in lethal outcome

According to the recommendations in guideline [7] , the death risk levels of coronary heart disease are categorized into three levels: low risk, intermediate risk, and high risk. The risk thresholds are set based on the model's probability density plots on the validation set. As shown in **Figure 3**(a), a probability below 0.2 indicates low risk (+), a probability between 0.2 and 0.6 suggests intermediate risk (++), and a probability beyond 0.6 indicates high risk (+++). These findings indicate a positive correlation between increasing risk levels and higher mortality rates, which can be found in **Figure 3**(b).

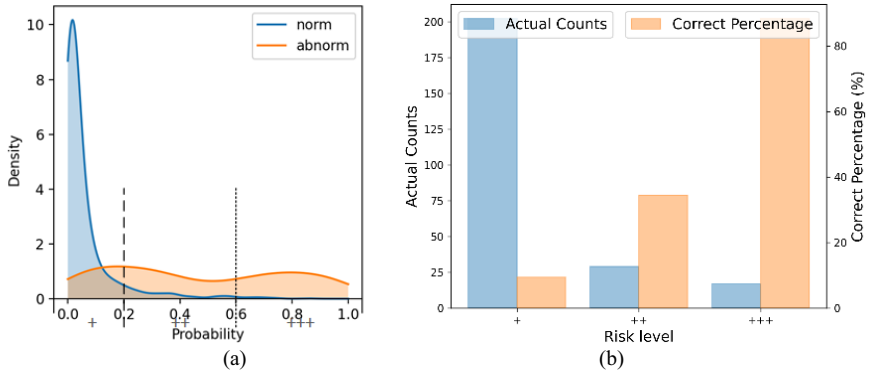


**Figure 3.** Lethal risk assessment by DeepFM model. (a) Probability density plots of training set. (b) Risk prediction results in test set.

## 5. Discussion

In addition to analyzing the prediction accuracy of classification models for different labels, it is also possible to further analyze the interpretability of the models. For deep learning methods, it's usually hard to find the balance between accuracy and interpretability. Here, the SHAP (SHapley Additive exPlanations) [8] method is utilized to explain the input features, which is congruent with human intuition. The SHAP method should fulfill the addictive attribute as shown in Equation (3).

$$f(x) = g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x_i' \tag{3}$$

The $g(x')$ signifies the explanation model that matches the original model $f(x)$ when $x = h_x(x')$, where $\phi_0 = f(h_x(0))$ denotes the base value.

As shown in Figure 4, SHAP values explain the contribution of each feature to the prediction of a given black box model. The bar charts of Figure 4(a) explain the impact of features on multi-label classification. Figure 4(b) depicts a beeswarm plot explaining the key feature influences for the lethal label. After inspection, the SHAP with NN model fulfills better the Equation (3), while the SHAP with RF model has larger error. Furthermore, we use force plots on individual cases to examine local interpretability of feature importance of the NN model, as shown in Figure 4(c). Taking the negative sample as an example, the base value of the model is 0.24, which signifies the average risk indicator and the threshold of MCC. The predictive value of the original model is 0.07, so this sample is documented as benign, and the feature AGE has negative feature contribution in this specific sample.

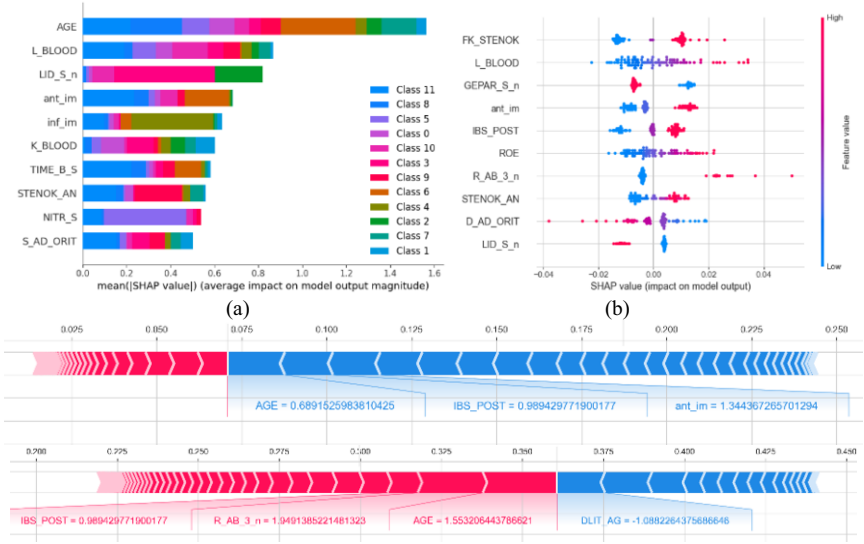


**Figure 4.** Model interpretability. (a) Top 10 SHAP value for multi-label classification. (b) Top 10 SHAP value for binary classification. (c) Local interpretability (negative sample for upper while positive sample for lower).

## 6. Conclusion

While traditional machine learning methods are commonly used for tabular data, this study aims to investigate the performance of various deep learning models in predicting the risk of Myocardial infarction complications. In terms of multi-label classification, some evaluation metrics of the NN model are surpassed by traditional machine learning methods, but overall performance is similar. In binary classification, the DeepFM model proposed in this study exhibits higher accuracy in predicting myocardial infarction samples with potentially lethal risks in this dataset. In addition, this study conducts interpretability analysis of the models, thus dispelling the opaque nature of deep learning models, thereby facilitating future research on the deep integration of ECG signals.

## Acknowledgements

This work was supported by China Postdoctoral Science Foundation (Grant No. 2023M732219) and Shanghai Municipal Science and Technology Major Project (No. 2021SHZDZX0102).

## References

[1] P. W. F. Wilson, R. B. D'Agostino, D. Levy, A. M. Belanger, H. Silbershatz, and W. B. Kannel, 'Prediction of Coronary Heart Disease Using Risk Factor Categories', *Circulation*, vol. 97, no. 18, pp. 1837–1847, May 1998, doi: 10.1161/01.CIR.97.18.1837.

[2] M. E. Charlson, P. Pompei, K. L. Ales, and C. R. MacKenzie, 'A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation', *Journal of Chronic Diseases*, vol. 40, no. 5, pp. 373–383, Jan. 1987, doi: 10.1016/0021-9681(87)90171-8.

[3] L. Grinsztajn, E. Oyallon, and G. Varoquaux, 'Why do tree-based models still outperform deep learning on typical tabular data?', in *36th Conference on Neural Information Processing Systems (NeurIPS 2022) Track on Datasets and Benchmarks*, in Advances in Neural Information Processing. New Orleans, United States, Nov. 2022. Accessed: Aug. 16, 2023. [Online]. Available: https://hal.science/hal-03723551

[4] M.-L. Zhang and Z.-H. Zhou, 'A Review on Multi-Label Learning Algorithms', *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1819–1837, Aug. 2014, doi: 10.1109/TKDE.2013.39.

[5] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, 'DeepFM: A Factorization-Machine based Neural Network for CTR Prediction', in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, Aug. 2017, pp. 1725–1731. doi: 10.24963/ijcai.2017/239.

[6] Golovenkin, S.E.; Gorban, Alexander; Mirkes, Evgeny; Shulman, V.A.; Rossiev, D.A.; Shesternya, P.A, 'Myocardial infarction complications Database'. University of Leicester, Jul. 07, 2020.

[7] M. Gulati *et al.*, '2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines', *Journal of the American College of Cardiology*, vol. 78, no. 22, pp. e187–e285, Nov. 2021, doi: 10.1016/j.jacc.2021.07.053.

[8] S. M. Lundberg and S.-I. Lee, 'A Unified Approach to Interpreting Model Predictions', in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. Accessed: Aug. 09, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html