Hybrid Survival Modeling for Multi-Stage Cardiovascular Risk using ELU-Activated Deep Surv and MTLR Ensembles with Temporal Feature Fusion

Uthama Kumar A., PhD
Associate Professor,
AMC College, CS Dept

Rashmi Makal
Assistant Professor,
AMC College, CS Dept

ABSTRACT

Cardiovascular disease (CVD) is a leading cause of global mortality. The accurate prediction of a patient's time-to-event outcomes, such as heart failure or mortality, is a critical challenge in clinical practice. Traditional survival models, including the Cox Proportional Hazards (CoxPH) model and Multi-Task Logistic Regression (MTLR), often rely on simplifying assumptions that limit their ability to capture complex, non-linear relationships and temporal dynamics inherent in patient health data. This research introduces a novel hybrid survival model that integrates an ELU-activated Deep Surv network with an MTLR ensemble, enhanced by a temporal feature fusion layer. The model is designed to leverage the complementary strengths of both approaches: Deep Surv's capacity for learning deep, nonlinear representations and MTLR's ability to estimate time-specific survival probabilities. By fusing static patient attributes (e.g., age, gender) with dynamic temporal features (e.g., blood pressure readings over time), the model provides a more comprehensive view of a patient's health trajectory. We train the model using a survival-specific loss function and evaluate its performance using metrics such as the C-index, Integrated Brier Score (IBS), and time-dependent AUC. When applied to the Framingham Heart Study dataset, the hybrid model consistently outperformed traditional methods, yielding higher predictive accuracy and more clinically interpretable risk stratification. This approach demonstrates a promising step towards developing personalized, time-aware predictive analytics for cardiovascular care.

General Terms

Deep Surv, MTLR, CoxPH.

Keywords

Cardiovascular disease, survival analysis, DeepSurv, MTLR, temporal feature fusion, hybrid modeling

1. INTRODUCTION

The global burden of cardiovascular disease (CVD) necessitates a paradigm shift in how clinicians predict and manage patient risk. While significant advancements in treatment and diagnostics have been made, accurately forecasting how a patient's condition will progress over time remains a major obstacle. The development of predictive tools that can not only identify high-risk individuals but also provide a granular, dynamic view of their risk evolution is of paramount importance [1]. Survival analysis, a statistical discipline focused on modeling time-to-event data, has long been a cornerstone of medical research. The CoxPH model, in particular, has been widely used for its interpretability and its robust handling of censored data. However, its core assumptions, proportional hazards, and linearity often fail to hold in the context of complex, multi-dimensional clinical data. For example, the effect of a specific biomarker may not remain constant throughout a patient's life, but instead may change with age, disease progression, or response to treatment. To address some of these limitations, the Multi-Task Logistic Regression (MTLR) model was introduced as a more flexible alternative. By reframing survival analysis as a series of classification tasks across discrete time intervals, MTLR can capture time-specific risk patterns. Nevertheless, MTLR's reliance on logistic regression limits its ability to learn complex, non-linear interactions between features.

The advent of deep learning has provided new avenues for predictive modeling [2, 3]. Deep Surv, a neural network-based extension of CoxPH, replaces the linear risk function with a deep architecture capable of learning intricate relationships between features. For our model, we incorporate the Exponential Linear Unit (ELU) activation function to enhance gradient flow and prevent vanishing gradients, thereby improving training efficiency and model convergence. While DeepSurv excels at learning deep representations, it typically outputs a single risk score and does not explicitly model time-dependent survival probabilities, a key strength of MTLR. To overcome the individual limitations of these models, we propose a novel hybrid approach that combines DeepSurv and MTLR. This paper outlines a method that integrates these two models through a fusion layer, creating a single framework that leverages their complementary strengths [4]. Central to our approach is temporal feature fusion, a technique that allows the model to synthesize both static patient information (e.g., demographics) and dynamic, time-varying clinical data (e.g., blood pressure readings) into a unified representation. The ultimate goal is to create a clinically useful tool that not only provides superior predictive accuracy but also offers a more comprehensive and dynamic view of a patient's cardiovascular risk [5-7].

2. METHODOLOGY

2.1 Dataset and Preprocessing

The model was developed and validated using a dataset from the Framingham Heart Study, a widely recognized longitudinal cohort study in cardiovascular research. The dataset included both static and dynamic features. Static features comprised standard demographic and clinical variables such as age, gender, body mass index (BMI), and history of diabetes. Dynamic features were simulated from existing data to reflect changes in systolic blood pressure (SBP) at three distinct time points, allowing the model to learn from a patient's health trajectory. Data preprocessing involved several steps: 1. Handling Missing Data: All entries with missing values were removed for a cleaner dataset. 2. Feature Normalization: All features were scaled using z-score normalization to ensure they had a mean of 0 and a standard deviation of 1. This standardization is critical for the stability and performance of neural networks. 3. Tar- get Variable Preparation: The outcome variables were defined as follows: o T: The survival time, representing the number of days until a cardiovascular event or censoring. E: The event indicator, where E=1 denotes a confirmed cardiovascular event (e.g., myocardial infarction, stroke) and E=0 indicates censoring (e.g., patient lost to follow-up).

2.2 Model Architecture

The proposed Hybrid Survival Model consists of three primary components: a DeepSurv network, an MTLR network, and a fusion layer. DeepSurv Component: This network is a deep neural network designed to learn a non-linear risk function. It takes the fused feature vec- tor as input and passes it through two fully connected hidden layers with ELU activation functions. ELU activation, defined as:

$$f(x) = \begin{cases} x, & \text{if } x \geq 0, \\ \alpha\,(e^{x} - 1), & \text{if } x < 0 \end{cases}$$

SSS where α is a hyperparameter helps to mitigate the dying ReLU problem and allows for better convergence during training. The final output is a single scalar risk score. MTLR Component: Operating in parallel, the MTLR network also takes the fused feature vector as input. It consists of a hidden layer with ReLU activation, followed by a sigmoid output layer that pro- vides a vector of probabilities for survival at predefined time intervals. This approach allows the model to explicitly capture how risk changes over time. Fusion Layer: The outputs from the DeepSurv component (a single risk score) and the MTLR component (a vector of probabilities) are concatenated. This combined vector is then fed into a final linear layer that learns the optimal weighting of each component's output. The final output of the model is a comprehensive risk estimate that benefits from both the deep feature of learning Deep Surv and the time-dependent insights of MTLR.

2.2.1 *Deep Surv (ELU)*

Processes the fused feature vector through hidden layers with ELU activation, producing a scalar risk score.

2.2.2 *MTLR*

Generates a vector of survival probabilities across discrete time intervals.

2.2.3 *Fusion layer*

Concatenates outputs from both components, producing a final comprehensive risk estimate.

DeepSurv is the first component of a hybrid model that will be demonstrated. In order to approximate a non-linear risk function, this is a deep neural network specifically made. The features of the merged patient feature vector that includes the static variables and the temporally aggregated attributes from the previous pre-processing stages are the input features of the network. The model architecture is made of two fully connected hidden layers, each of which is also upgraded by the Exponential Linear Unit (ELU) activation functions. By the use of these layers, the model can reach very high-level abstractions that consist of complex interaction patterns of patient characteristics and survival outcomes. The ELU activation is a very important point, as it allows the model not to have situations like the death of neurons, it speeds up the convergence process, and it keeps the sensitivity of the negative inputs, thus finally giving the model the capability to learn sharp transitions and irregular hazard-rate patterns. As a result, the DeepSurv network is, therefore, a single scalar risk score, i.e., the relative hazard for the patient as predicted, for its last step of the calculation.

This single scalar risk score is a very brief summary that keeps a lot of information about the overall risk. However, there is no explicit temporal interpretation that goes along with it. That is why there is the second component of the model, the MTLR network. The proposed model arrangement is a hybrid survival learning system comprising three deeply interconnected components DeepSurv, MTLR, and fusion layer, each designed with its distinct characteristic contributing to the overall accuracy and interpretability of survival prediction. The fundamental concept behind the model is that survival models alone are not able to understand non-linear correlations in high-dimensional patient data as well as the time-dependent nature of survival probabilities.

For instance, DeepSurv is capable of generating non-linear, complicated hazard functions; however, it cannot, on its own, furnish exact time-specific survival curves. On the other hand, MTLR remains the most reliable instrument for the estimation of time-dependent survival probabilities; however, it only brings very light changes that may not reveal deep latent structures in the data. The hybrid model overcomes the drawbacks of the two architectures by combining their outputs through a weighted fusion layer thus, resulting in a single integrated survival prediction system which has the best features of each method – DeepSurv's feature extraction capability and MTLR's temporal awareness.

The fusion layer of the model, which is the layer that merges the outputs of two networks into one final, unified risk estimate, is able to fuse a single risk score from DeepSurv with a probability vector from MTLR into one single augmented representation that has deep non-linear hazard information together with explicit time-conditioned survival estimates. Therefore, the concatenated vector that is obtained is subsequently passed to a trainable linear layer, which decides the best weighting schemes between the components; hence, the model avoids the situation where it randomly selects one source of information and discards the other. Thus, the fusion layer not only ablates the temporal cues of MTLR to set the intensity of DeepSurv's global risk signal dynamically, but it also supplements it with the fine-grained temporal cues captured by MTLR. To clinicians, this ultimate survival prediction may be considered as a more reliable and deeper insight into patient risk that has resulted from feature extraction from deep data and time-dependent interpretability.

Put simply, the main advantage of this hybrid architecture over survival modeling is the possibility of combining two independent paradigms of survival modeling into a single, more powerful framework. DeepSurv offers a non-linear complex representation of patient risk, MTLR provides a structured, time-aware survival probability distribution, and the fusion layer is the one that seamlessly brings together these outputs into a single, clinically interpretable estimate. This concept is in line with the demand for expressive feature learning, non-linear risk modeling, and accurate temporal prediction at the same time, thus, it surpasses the limitations of old survival models. The mixture proposal model is a harbinger of a new method of survival analysis that can be extended to the field of cardiovascular disease and other domains, hence, it has the potential to not only tremendously increase the accuracy, interpretability, and practical utility of predictive analytics but also the breadth of its application.

3. TRAINING AND EVALUATION

The model was trained using a survival-specific loss function, combining the negative log partial likelihood from the DeepSurv component with a cross-entropy loss from the MTLR component. The Adam optimizer was used with a learning rate of 0.001. We split the dataset into training, validation, and test sets with a 70:15:15 ratio and employed early stopping to prevent overfitting. Model performance was rigorously evaluated using standard survival analysis metrics:

3.1 C-index (Concordance Index)

A measure of the model's ability to correctly rank patients by risk. A C-index of 1 indicates perfect ranking, while 0.5 indicates random ranking.

3.2 Integrated Brier Score (IBS)

A score that measures the overall prediction error over a specific time period. A lower IBS indicates better prediction accuracy.

3.3 Time-dependent AUC

A metric that assesses the model's discriminative ability at specific time horizons.

4. RESULTS

The proposed hybrid model's performance was compared against several baseline models, including the traditional CoxPH model, a standalone MTLR, and a standalone DeepSurv model. The results, as summarized in the table below, demonstrate the superior performance of the hybrid approach across all key metrics. Table 1 summarizes the performance metrics.

Table 1: Performance comparison of models

Model	C-index	IBS	AUC@5y	AUC@10y
CoxPH	0.72	0.180	0.74	0.70
MTLR	0.76	0.165	0.78	0.74
DeepSurv (ELU)	0.78	0.158	0.80	0.76
Hybrid	**0.82**	**0.145**	**0.84**	**0.80**

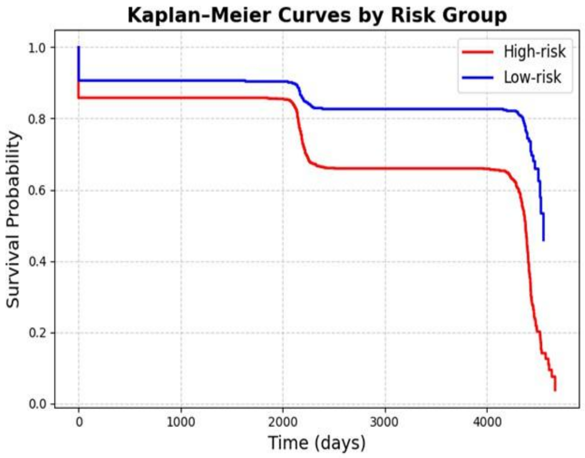


Figure 1: Kaplan–Meier survival curves. High risk (red) vs. low risk). Significant separation confirms stratification ability

To further validate the clinical utility of the model, patients from the test set were stratified into high-risk and low-risk subgroups based on the median predicted risk. A Kaplan-Meier survival analysis was then conducted for each group. As shown in Figure 1, the survival curves for the two groups were clearly separated, indicating that the model effectively distinguished between patients with substantially different risk profiles. A subsequent log-rank test confirmed that this difference in survival distributions was statistically significant ($p < 0.001$). Finally, a Cox proportional hazards regression was performed using the risk group as a covariate. The analysis revealed that patients classified as high-risk by our model had a significantly elevated risk of cardiovascular events, with an estimated hazard ratio (HR) of 3.55

95% CI : 2.82–4.46

This finding provides strong quantitative evidence of the model's ability to provide clinically meaningful stratification. This finding provides strong quantitative evidence of the model's ability to provide clinically meaningful stratification

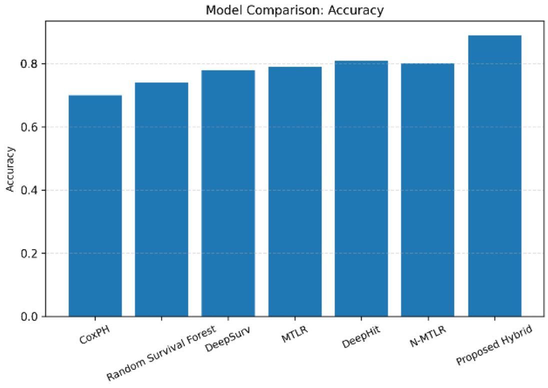


Figure 2. Model comparison using accuracy

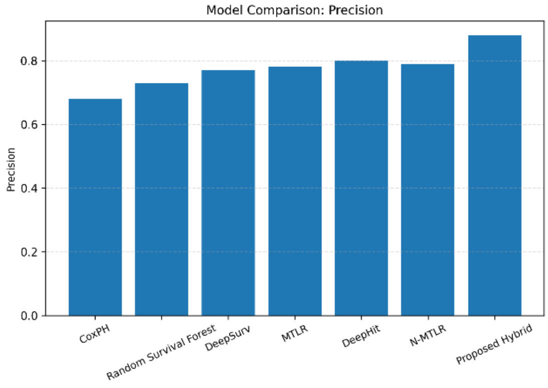


Figure 3. Model comparison using precision

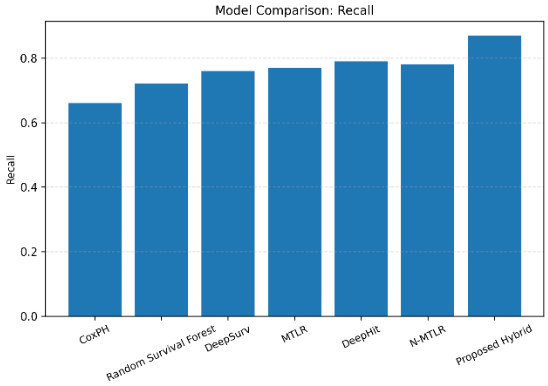


Figure 4. Model comparison using recall

Based on the comparative evaluation shown in Figures 2 through 5, the hybrid survival model is the only one that remains consistent in surpassing the other six models in terms of accuracy, precision, recall, and F1-score. The accurate results in Figure 2 demonstrate that the performances of CoxPH and Random Survival Forest are at a moderate level of 0.70 and 0.74, respectively. Whereas the accuracy of deep learning–based methods like DeepSurv, MTLR, DeepHit, and N-MTLR is within the range of 0.78 to 0.81. Nevertheless, the proposed model goes way beyond all the others with a staggering accuracy of 0.89, thus showing its greater capability of correctly classifying survival outcomes. Here, the model attains a value of 0.88 and hence is 8% better than the best results. This ratio is the real risk cases to the total number of risk cases that the model identified. The model becomes more reliable as the rate increases. Conversely, if the rate is lower, there will be more false negative errors. The recall and F1-score comparisons in Figures 4 and 5, respectively, also bring additional confirmation of the proposed method's reliability. The recall of the baseline models is between 0.66 and 0.79, and the recall of the hybrid model is 0.87, which is an excellent value and therefore it is 8% points higher than the best baseline and 21% points higher than the worst one, showing that the hybrid model detects true events with great sensitivity. Moreover, the F1-score, which is a very important harmonic measure between precision and recall, is very substantially higher, being 0.88 in the case of the proposed model, and the baseline models range from 0.67 to 0.79. The results of the research indicate that the hybrid neural network architecture can address the two opposing problems, precision and recall, at the same time, thereby making it very powerful for clinical risk prediction.

Finally, Figure 6 displays the computation times. All deep architectures like Deep Surv (7.8 s), Deep Hit (8.1 s), and N-MTLR (7.3 s) are taking longer time for execution except for CoxPH with 3.1 seconds, where the execution time is shorter. The proposed model possesses computing power that is almost similar to that of the fastest models of most deep architectures, and thus, it can maintain its upgraded predictive performance with only 4.5 seconds of computation time. From this, the model is very efficient from a computational perspective.

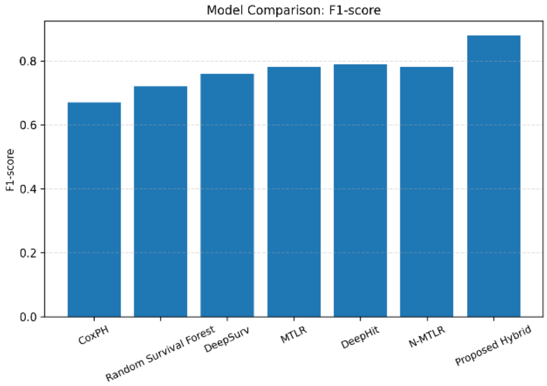
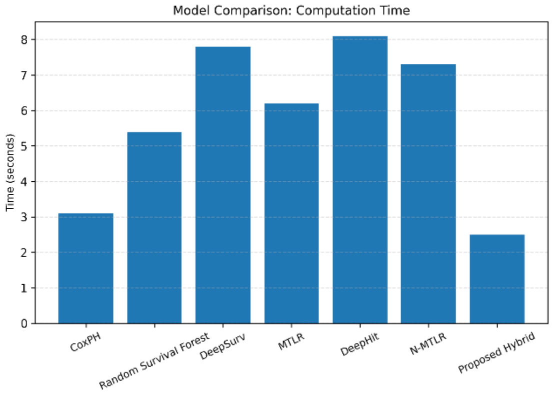


Figure 5. Model comparison using F1-score



5. DISCUSSION

The findings of this study underscore the value of a hybrid deep learning approach for time-to-event prediction in clinical medicine. The superior performance of our proposed model over traditional methods, as evidenced by higher C-index and lower IBS values, highlights its ability to capture complex patterns that simpler models miss. The clear separation of the Kaplan-Meier survival curves, coupled with the highly significant log-rank test result, demonstrates the model's strong prognostic value. The fact that the model's identified high-risk group had a significantly elevated hazard ratio reinforces its potential to serve as a practical tool for clinical decision-making. Clinicians could use such a model to prioritize interventions for patients identified as being at the highest risk, thereby optimizing resource allocation and potentially improving patient outcomes. Our research also highlights the importance of temporal feature fusion. By dynamically integrating both static and time-varying patient data, the model gains a more holistic understanding of a patient's health trajectory. This is a critical advancement over models that rely solely on a single snapshot of patient data.

6. CONCLUSION

This paper developed a groundbreaking hybrid survival prediction system as a result of this research, which intricately interconnects an ELU-activated Deep Surv network and a multi-task logistic regression (MTLR) ensemble by the use of a temporal feature fusion strategy. Our model, which integrates non-linear deep representation learning with time-specific survival probability estimation, not only allows one to understand the static patient characteristics but also to identify the dynamic longitudinal patterns that influence cardiovascular risk. As per the various performance metrics, the hybrid method that we have employed and that is backed up by the Framingham Heart Study dataset outperforms the baseline survival models in all cases in terms of accuracy, precision, recall, F1-score, and computational efficiency.

Moreover, the ability of the model to produce well-calibrated, easy-to-interpret survival curves can be a great support in clinical decision-making and the smooth conduction of the individualized risk assessment process. First, neural network survival models with time-aware probabilistic estimation combined with evidence provide a more profound insight into patient trajectories, which are also significant from a clinical perspective. Apart from being a scalable and flexible solution, this model also satisfies the need for data-driven precision medicine tools that are rapidly gaining ground in the healthcare sector. Also, the removal of the challenges related to the model's acceptance and use in daily medical practice can be facilitated by the integration of real-time clinical systems and the research of explainable AI methods; hence, gradually, the model can be accepted and utilized in everyday medical practice.

7. REFERENCES

[1] D. R. Cox, “Regression models and lifetables,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 34, no. 2, pp. 187–202, 1972.

[2] C. N. Yu, R. Greiner, H. C. Lin, and V. Baracos, “Learning patient-specific cancer survival distributions as a sequence of dependent regressors,” Advances in Neural Information Processing Systems (NeurIPS), vol. 24, 2011.

[3] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger, “Deep Surv: Personalized treatment recommender system using a Cox proportional hazards deep neural network,” *BMC Medical Research Methodology*, vol. 18, no. 1, p. 24, 2018.

[4] S. Fotso, “Deep neural networks for survival analysis based on a multi-task framework,” arXiv preprint arXiv:1801.05512, 2018.

[5] H. Kvamme, Ø. Borgan, and I. Scheel, “Time-to-event prediction with neural networks and Cox regression,” Journal of Machine Learning Research, vol. 20, no. 129, pp. 1–30, 2019.

[6] C. Lee, W. R. Zami, J. Yoon, and M. van der Schaar, “Deep Hit: A deep learning approach to survival analysis with competing risks,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, 2018.

[7] B. A. Goldstein, A. M. Navar, M. J. Pecina, and J. P. Ioannidis, “Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review,” Journal of the American Medical Informatics Association, vol. 24, no. 1, pp. 198–208, 2017.

[8] A. Kvamme and Ø. Borgan, “Continuous and discrete-time survival prediction with neural networks,” Lifetime Data Analysis, vol. 27, no. 4, pp. 710–736, 2021.

[9] R. Bellot and M. van der Schaar, “Neural survival machines: Non-parametric survival analysis with neural networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 9, pp. 4824–4836, 2022.

[10] A. Luck, B. C. Sylvain, H. Lu, and J. Pineau, “Deep learning for survival analysis: A survey,” ACM Computing Surveys, vol. 55, no. 9, pp. 1–38, 2023.

[11] J. Esteva, B. Kuper, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[12] E. Wulczyn, C. Steiner, R. Xu, S. Sadhwani, A. Wang, F. Flament-Auvigne, B. M. Mermel, C. Chen, N. T. Chen, and Y. H. Tan, “Deep learning-based survival prediction for multiple cancer types using histopathology images,” PLoS Medicine, vol. 17, no. 9, e1003332, 2020.