iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** II **Month of publication:** February 2024

DOI: https://doi.org/10.22214/ijraset.2024.58349

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

An Explainable Transformer-Based Deep Learning Model for the Prediction of Incident Heart Failure

Mohammed Ejaz[1], Dr. M. Arathi[2]
[1]*Student, M.Tech (Data Sciences), Department of Information Technology, JNTU Hyderabad, Hyderabad, India*
[2]*Professor of CSE, Department of Information Technology, JNTU Hyderabad, Hyderabad, India*

Abstract: ***Predicting the incidence of complex chronic conditions such as heart failure is challenging. Deep learning models applied to rich electronic health records may improve prediction but remain unexplainable hampering their wider use in medical practice. We aimed to develop a deep-learning framework for accurate and yet explainable prediction of 6-month incident heart failure (HF). Using 100,071 patients from longitudinal linked electronic health records. we applied a novel Transformer based risk model using all community and hospital diagnoses and medications contextualized within the age and calendar year for each patient's clinical encounter. Feature importance was investigated with an ablation analysis to compare model performance when alternatively removing features and by comparing the variability of temporal representations. A post-hoc perturbation technique was conducted to propagate the changes in the input to the outcome for feature contribution analyses. Our model achieved 0.93 area under the receiver operator curve and 0.69 area under the precision-recall curve on internal 5-fold cross validation and outperformed existing deep learning models. Ablation analysis indicated medication is important for predicting HF risk, calendar year is more important than chronological age, which was further reinforced by temporal variability analysis. Contribution analyses identified risk factors that are closely related to HF. Many of them were consistent with existing knowledge from clinical and epidemiological research but several new associations were revealed which had not been considered in expert driven risk prediction models. In conclusion, the results highlight that our deep learning model, in addition high predictive performance, can inform data-driven risk factor identification.***

Keywords: BEHRT Transformer; Deep Neural Network; Healthcare Data Analytics; Exploratory Data Analysis.

I. INTRODUCTION

Heart failure (HF) remains a major cause of morbidity, mortality, and economic burden. Despite recent evidence suggesting improvements in the quality of clinical care that patients with HF receive, and favorable trends in prognosis, the incidence of HF has changed little. Indeed, as a consequence of population growth and ageing, the absolute burden of HF has been increasing, with incidence rates similar to the four most common causes of cancer combined. These observations reinforce the need for fuller implementation of existing strategies for HF prevention and further investigations into risk factors. Several statistical models have been developed to predict risk of incident HF; however, the predictive performance of these models has been largely unsatisfactory. The growing availability of comprehensive clinical datasets, such as linked electronic health records (EHR) with extensive clinical information from a large number of individuals, together with advances in machine learning, offer new opportunities for developing more robust risk-prediction models than conventional statistical approaches. Such data-driven approaches can also potentially discover new associations that are less dependent on expert knowledge. However, empirical evidence of robust prediction of complex chronic conditions, such as HF is limited. Prominent deep learning (DL) architectures have shown modest performance in large-scale, complex EHR datasets for risk prediction of various conditions including HF. Due to their high level of abstraction, these DL models have typically had poor "explainability" or ability to demonstrate results in a language understandable by humans. This has limited their trustfulness and contribution to risk factor discoveries and wider clinical adoption. Recent research has shown progress in explaining DL models in the fields of natural language and computer vision and methods such as saliency map and feature perturbation have gained wide popularity. However, explainable DL with rich EHR is still in its nascency; hence, tailoring known methods to improve model explainability in the medical context is crucial.

II. LITERATURE SURVEY

1) Risk prediction models for incident heart failure: A systematic review of methodology and model performance.

Numerous models predicting the risk of incident heart failure (HF) have been developed; however, evidence of their methodological rigor and reporting remains unclear.

©IJRASET: All Rights are Reserved

This study critically appraises the methods underpinning incident HF risk prediction models. Methods and results: EMBASE and PubMed were searched for articles published between 1990 and June 2016 that reported at least 1 multivariable model for prediction of HF. Model development information, including study design, variable coding, missing data, and predictor selection, was extracted. Nineteen studies reporting 40 risk prediction models were included. Existing models have acceptable discriminative ability (C-statistics > 0.70), although only 6 models were externally validated. Candidate variable selection was based on statistical significance from a univariate screening in 11 models, whereas it was unclear in 12 models. Continuous predictors were retained in 16 models, whereas it was unclear how continuous variables were handled in 16 models. Missing values were excluded in 19 of 23 models that reported missing data, and the number of events per variable was < 10 in 13 models. Only 2 models presented recommended regression equations. There was significant heterogeneity in discriminative ability of models with respect to age (P < .001) and sample size (P = .007). Conclusions: There is an abundance of HF risk prediction models that had sufficient discriminative ability, although few are externally validated. Methods not recommended for the conduct and reporting of risk prediction modeling were frequently used, and resulting algorithms should be applied with caution.

2) Temporal trends and patterns in mortality after incident heart failure a longitudinal analysis of 86 000 individuals.

Importance: Despite considerable improvements in heart failure care, mortality rates among patients in high-income countries have changed little since the early 2000s. Understanding the reasons underlying these trends may provide valuable clues for developing more targeted therapies and public health strategies. Objective: To investigate mortality rates following a new diagnosis of heart failure and examine changes over time and by cause of death and important patient features. Design, setting, and participants: This population-based retrospective cohort study analyzed anonymized electronic health records of individuals who received a new diagnosis of heart failure between January 2002 and December 2013 who were followed up until December 2014 from the Clinical Practice Research Datalink, which links information from primary care, secondary care, and the national death registry from a subset of the UK population. The data were analyzed from January 2018 to February 2019. Main outcomes and measures: All-cause and cause-specific mortality rates at 1 year following diagnosis. Poisson regression models were used to calculate rate ratios (RRs) and 95% confidence intervals comparing 2013 with 2002, adjusting for age, sex, region, socioeconomic status, and 17 major comorbidities. Results: Of 86 833 participants, 42 581 (49%) were women, 51 215 (88%) were white, and the mean (SD) age was 76.6 (12.6) years. While all-cause mortality rates declined only modestly over time (RR comparing 2013 with 2002, 0.94; 95% CI, 0.88-1.00), underlying patterns presented explicit trends. A decline in cardiovascular mortality (RR, 0.73; 95% CI, 0.67-0.80) was offset by an increase in noncardiovascular deaths (RR, 1.22; 95% CI, 1.11-1.33). Subgroup analyses further showed that overall mortality rates declined among patients younger than 80 years (RR, 0.79; 95% CI, 0.71-0.88) but not among those older than 80 years (RR, 0.97; 95% CI, 0.90-1.06). After cardiovascular causes (898 [43%]), the major causes of death in 2013 were neoplasms (311 [15%]), respiratory conditions (243 [12%]), and infections (13%), the latter 2 explaining most of the observed increase in noncardiovascular mortality. Conclusions and relevance: Among patients with a new heart failure diagnosis, considerable progress has been achieved in reducing mortality in young and middle-aged patients and cardiovascular mortality across all age groups. Improvements to overall mortality are hindered by high and increasing rates of noncardiovascular events. These findings challenge current research priorities and management strategies and call for a greater emphasis on associated comorbidities. Specifically, infection prevention presents as a major opportunity to improve prognosis.

3) Temporal trends and patterns in heart failure incidence: A population-based study of 4 million individuals.

Large-scale and contemporary population-based studies of heart failure incidence are needed to inform resource planning and research prioritization but current evidence is scarce. We aimed to assess temporal trends in incidence and prevalence of heart failure in a large general population cohort from the UK, between 2002 and 2014. Methods: For this population-based study, we used linked primary and secondary electronic health records of 4 million individuals from the Clinical Practice Research Datalink (CPRD), a cohort that is representative of the UK population in terms of age and sex. Eligible patients were aged 16 years and older, had contributed data between Jan 1, 2002, and Dec 31, 2014, had an acceptable record according to CPRD quality control, were approved for CPRD and Hospital Episodes Statistics linkage, and were registered with their general practice for at least 12 months. For patients with incident heart failure, we extracted the most recent measurement of baseline characteristics (within 2 years of diagnosis) from electronic health records, as well as information about comorbidities, socioeconomic status, ethnicity, and region. We calculated standardized rates by applying direct age and sex standardization to the 2013 European Standard Population, and we inferred crude rates by applying year-specific, age-specific, and sex-specific incidence to UK census mid-year population estimates.

We assumed no heart failure for patients aged 15 years or younger and report total incidence and prevalence for all ages (>0 years). Findings: From 2002 to 2014, heart failure incidence (standardized by age and sex) decreased, similarly for men and women, by 7% (from 358 to 332 per 100 000 person-years; adjusted incidence ratio 0·93, 95% CI 0·91-0·94). However, the estimated absolute number of individuals with newly diagnosed heart failure in the UK increased by 12% (from 170 727 in 2002 to 190 798 in 2014), largely due to an increase in population size and age. The estimated absolute number of prevalent heart failure cases in the UK increased even more, by 23% (from 750 127 to 920 616). Over the study period, patient age and multi-morbidity at first presentation of heart failure increased (mean age 76·5 years [SD 12·0] to 77·0 years [12·9], adjusted difference 0·79 years, 95% CI 0·37-1·20; mean number of comorbidities 3·4 [SD 1·9] vs 5·4 [2·5]; adjusted difference 2·0, 95% CI 1·9-2·1). Socioeconomically deprived individuals were more likely to develop heart failure than were affluent individuals (incidence rate ratio 1·61, 95% CI 1·58-1·64), and did so earlier in life than those from the most affluent group (adjusted difference -3·51 years, 95% CI -3·77 to -3·25). From 2002 to 2014, the socioeconomic gradient in age at first presentation with heart failure widened. Socioeconomically deprived individuals also had more comorbidities, despite their younger age. Interpretation: Despite a moderate decline in standardized incidence of heart failure, the burden of heart failure in the UK is increasing, and is now similar to the four most common causes of cancer combined. The observed socioeconomic disparities in disease incidence and age at onset within the same nation point to a potentially preventable nature of heart failure that still needs to be tackled. Funding: British Heart Foundation and National Institute for Health Research.

4) *Predicting the risk of emergency admission with machine learning: Development and validation using linked electronic health records.*

Emergency admissions are a major source of healthcare spending. We aimed to derive, validate, and compare conventional and machine learning models for prediction of the first emergency admission. Machine learning methods are capable of capturing complex interactions that are likely to be present when predicting less specific outcomes, such as this one. Methods and findings: We used longitudinal data from linked electronic health records of 4.6 million patients aged 18-100 years from 389 practices across England between 1985 to 2015. The population was divided into a derivation cohort (80%, 3.75 million patients from 300 general practices) and a validation cohort (20%, 0.88 million patients from 89 general practices) from geographically distinct regions with different risk levels. We first replicated a previously reported Cox proportional hazards (CPH) model for prediction of the risk of the first emergency admission up to 24 months after baseline. This reference model was then compared with 2 machine learning models, random forest (RF) and gradient boosting classifier (GBC). The initial set of predictors for all models included 43 variables, including patient demographics, lifestyle factors, laboratory tests, currently prescribed medications, selected morbidities, and previous emergency admissions. We then added 13 more variables (marital status, prior general practice visits, and 11 additional morbidities), and also enriched all variables by incorporating temporal information whenever possible (e.g., time since first diagnosis). We also varied the prediction windows to 12, 36, 48, and 60 months after baseline and compared model performances. For internal validation, we used 5-fold cross-validation. When the initial set of variables was used, GBC outperformed RF and CPH, with an area under the receiver operating characteristic curve (AUC) of 0.779 (95% CI 0.777, 0.781), compared to 0.752 (95% CI 0.751, 0.753) and 0.740 (95% CI 0.739, 0.741), respectively. In external validation, we observed an AUC of 0.796, 0.736, and 0.736 for GBC, RF, and CPH, respectively.

The addition of temporal information improved AUC across all models. In internal validation, the AUC rose to 0.848 (95% CI 0.847, 0.849), 0.825 (95% CI 0.824, 0.826), and 0.805 (95% CI 0.804, 0.806) for GBC, RF, and CPH, respectively, while the AUC in external validation rose to 0.826, 0.810, and 0.788, respectively.

This enhancement also resulted in robust predictions for longer time horizons, with AUC values remaining at similar levels across all models. Overall, compared to the baseline reference CPH model, the final GBC model showed a 10.8% higher AUC (0.848 compared to 0.740) for prediction of risk of emergency admission within 24 months. GBC also showed the best calibration throughout the risk spectrum. Despite the wide range of variables included in models, our study was still limited by the number of variables included; inclusion of more variables could have further improved model performances. Conclusions: The use of machine learning and addition of temporal information led to substantially improved discrimination and calibration for predicting the risk of emergency admission. Model performance remained stable across a range of prediction time windows and when externally validated.

These findings support the potential of incorporating machine learning models into electronic health records to inform care and service planning.

5) Artificial Intelligence in Cardiology

Artificial intelligence and machine learning are poised to influence nearly every aspect of the human condition, and cardiology is not an exception to this trend. This paper provides a guide for clinicians on relevant aspects of artificial intelligence and machine learning, reviews selected applications of these methods in cardiology to date and identifies how cardiovascular medicine could incorporate artificial intelligence in the future. In particular, the paper first reviews predictive modeling concepts relevant to cardiology such as feature selection and frequent pitfalls such as improper dichotomization. Second, it discusses common algorithms used in supervised learning and reviews selected applications in cardiology and related disciplines. Third, it describes the advent of deep learning and related methods collectively called unsupervised learning, provides contextual examples both in general medicine and in cardiovascular medicine, and then explains how these methods could be applied to enable precision cardiology and improve patient outcomes. (J Am Coll Cardiol 2018;71:2668–79) © 2018 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation.

III. PROPOSED METHODOLOGY

In this study, we aimed to develop and validate a model for predicting incident HF, leveraging state-of-the-art, deep-sequential architecture applied to temporal and multi-modal EHR. We developed a novel Transformer deep-learning model for more accurate and yet explainable prediction of incident heart failure involving 100,071 patients from longitudinal linked electronic health records. On internal 5-fold cross validation and held-out external validation, our model achieved 0.93 and 0.93 area under the receiver operator curve and 0.69 and 0.70 area under the precision-recall curve, respectively and outperformed existing deep learning models. Predictor groups included all community and hospital diagnoses and medications contextualized within the age and calendar year for each patient's clinical encounter. The importance of contextualized medical information was revealed in a number of sensitivity analyses, and our perturbation method provided a way of identifying factors contributing to risk.

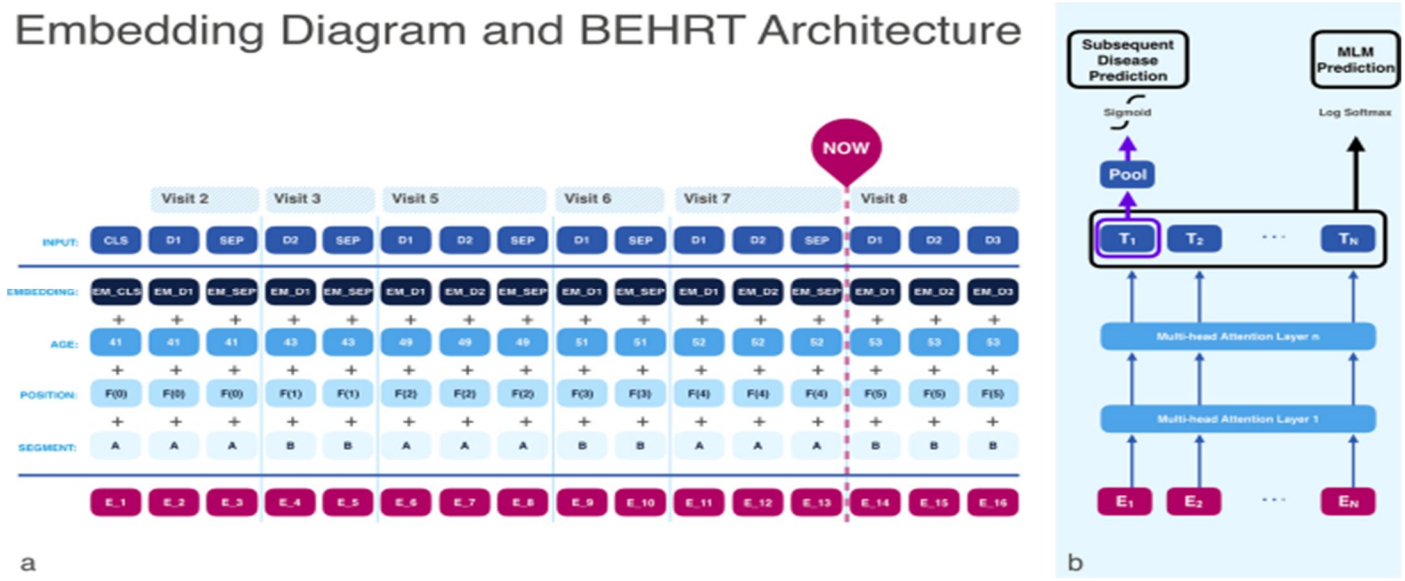


Fig. 1: BEHRT Embeddings and Architecture

IV. IMPLEMENTATION

1) *BEHRT*: Implement the explainable transformer-based deep learning model using TensorFlow or PyTorch.
2) *KNN (k-Nearest Neighbors)*: Train a k-Nearest Neighbors classifier on your dataset. KNN is a simple and interpretable algorithm that can serve as a baseline model.
3) *Logistic Regression*: Implement logistic regression, another interpretable model, and evaluate its performance.
4) *Decision Tree*: Build a decision tree classifier, which provides a clear decision-making path, making it interpretable.
5) *Random Forest*: Create a random forest classifier, an ensemble method that combines multiple decision trees for improved accuracy and interpretability.
6) *Voting Classifier (Logistic Regression + Random Forest + Gaussian Naive Bayes)*: Combine multiple classifiers using a voting mechanism. It allows you to take advantage of different classifiers' strengths and create a more robust and interpretable model.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

V. EXPERIMENTAL RESULTS

This section compared machine learning models applied to predict heart failure, It uses a the heat failure dataset. Firstly, the necessary libraries for data manipulation, visualization, modeling, and evaluation are imported to the program, Which are NumPy, Pandas, Matplotlib/Seaborn, Scikit-learn, and TensorFlow/PyTorch. Then, the heart failure dataset is loaded into a Pandas DataFrame for analysis.

In the exploratory data analysis phase, missing values are fixed using imputation or dropping, while various visualization techniques such as histograms, bar plots, and pair plots are visualized to understand data distribution and feature relationships. Also, numerical features are scaled using MinMaxScaler to standardize their range and give the model convergence during training.

After the data exploration, below given machine learning algorithms are implemented and evaluated. They are BEHRT model, k-Nearest Neighbors (KNN), logistic regression, decision tree, random forest, and a voting classifier that combines logistic regression, random forest, and Gaussian Naive Bayes models. Each algorithm is evaluated depending on its predictive accuracy and interpretability, this explains the suitability of different models for heart failure prediction tasks.

Below figures have the comparison of the models used and also the screenshots of the application. Comparison includes accuracy scores, precision, f1 and recall of each model.

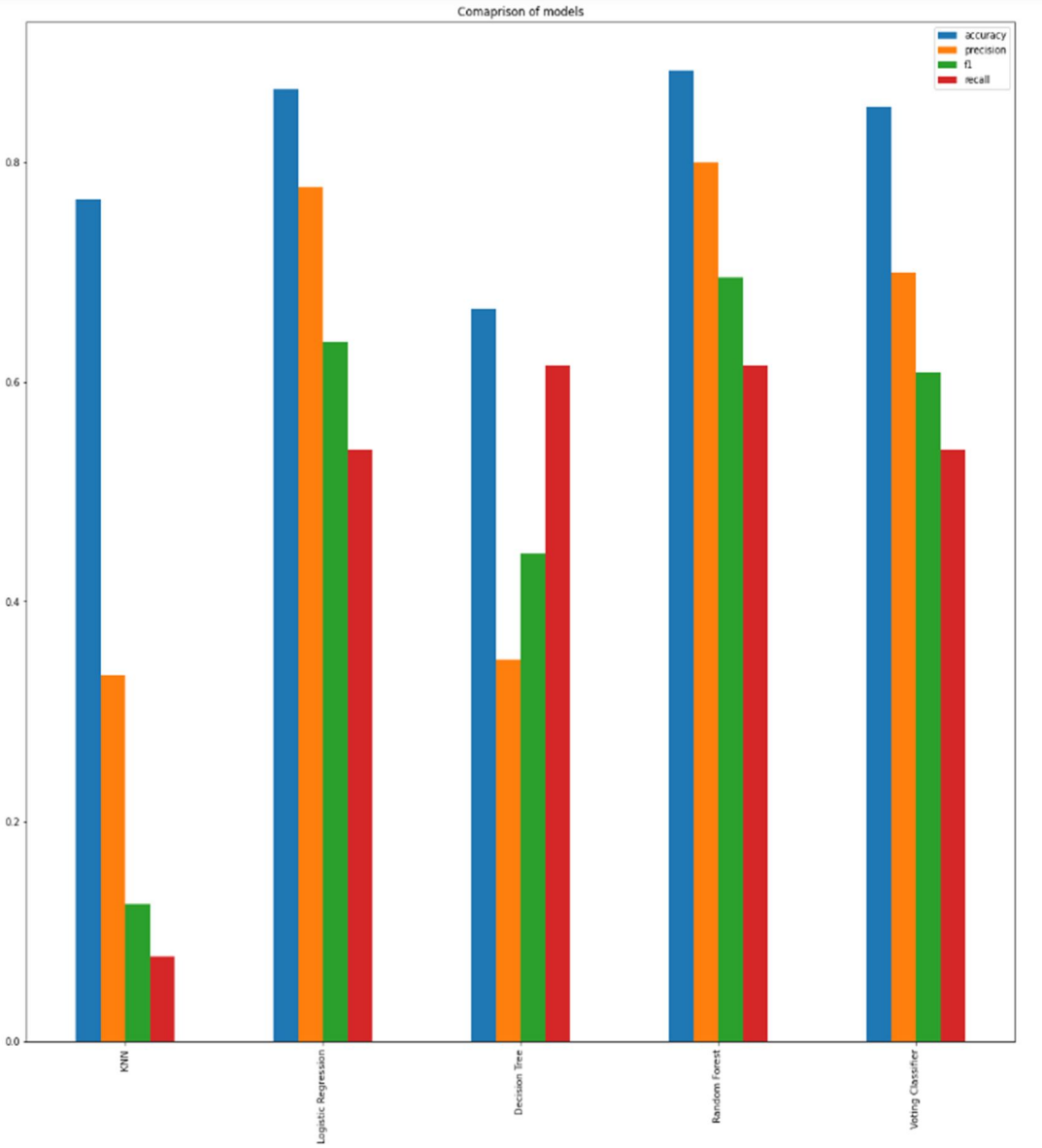


Fig. 2: Comparison of All the models KNN, Logistic Regression, Decision Tree, Random forest.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

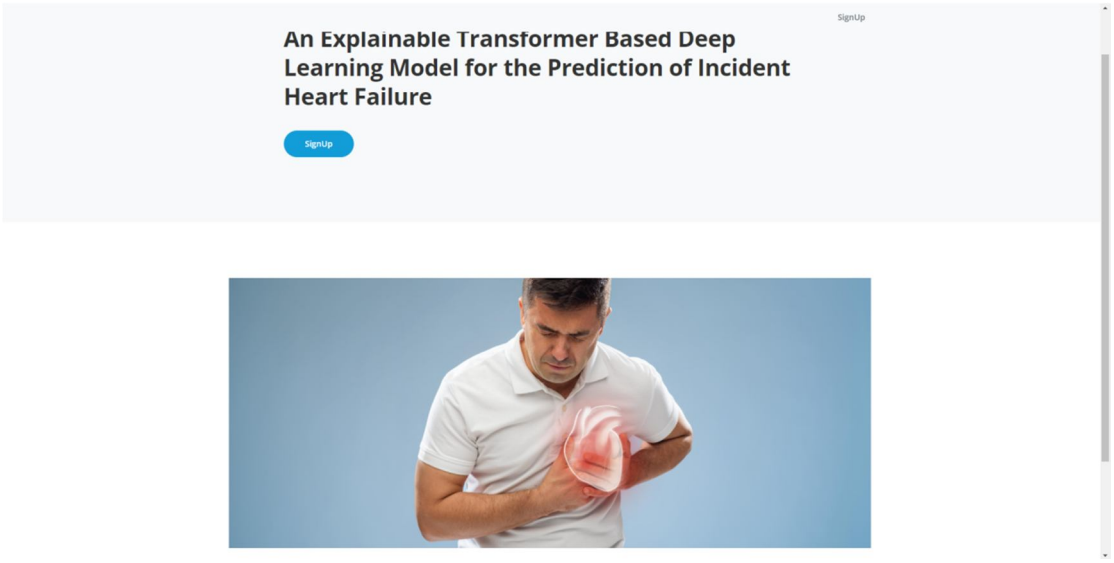


Fig. 3: Web App Sign up page.

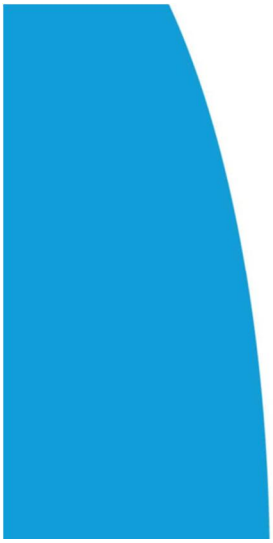


Fig. 4: Registration form

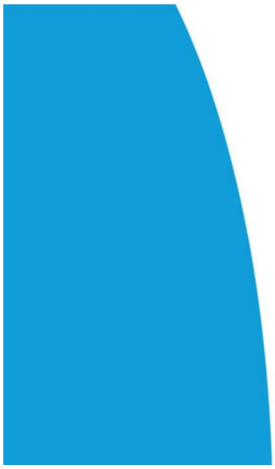
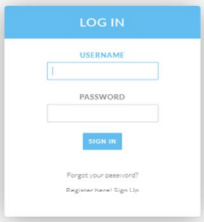



Fig. 5: Login Page

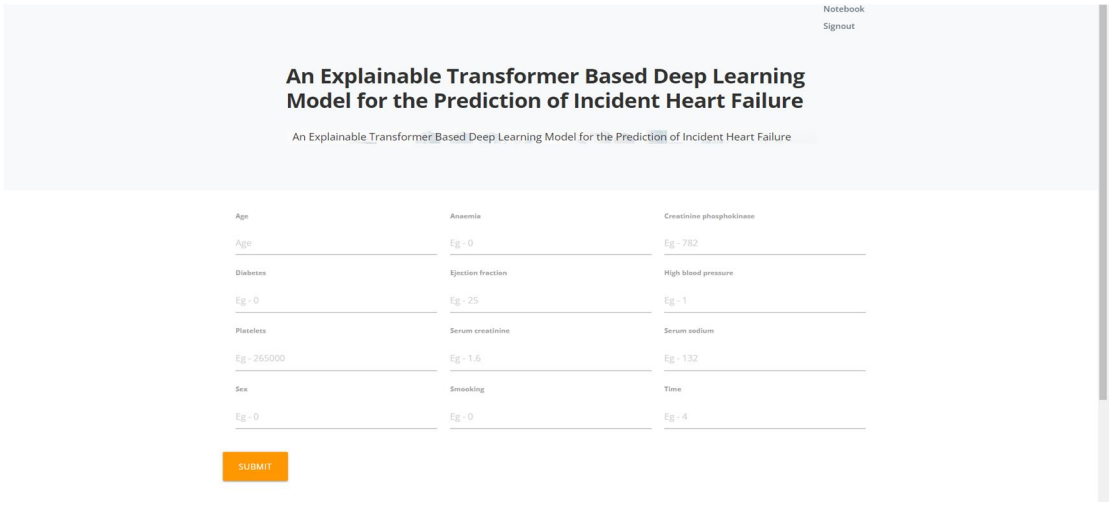


Fig. 6: Inputs Fields for the Model

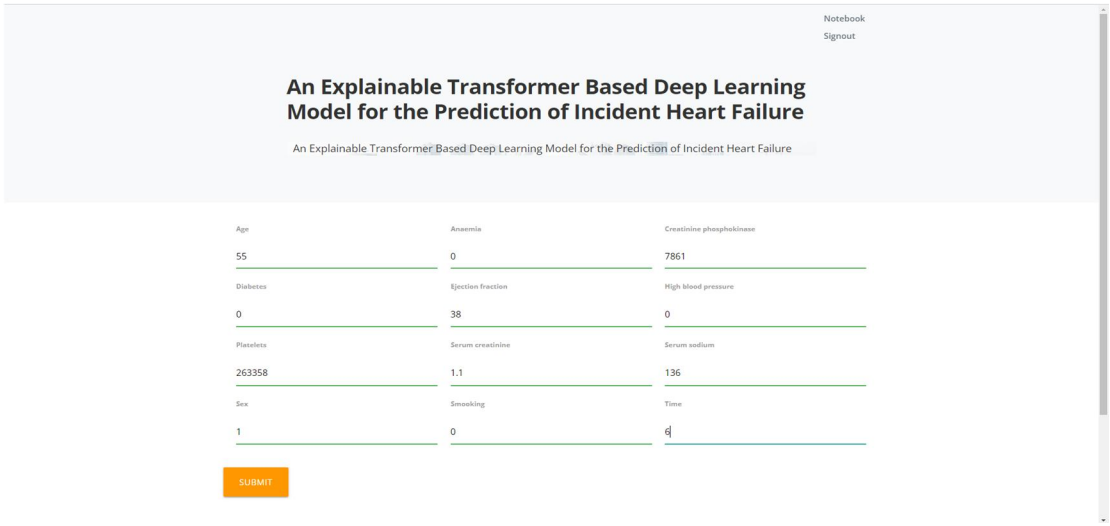


Fig 7: Adding Inputs to the Model

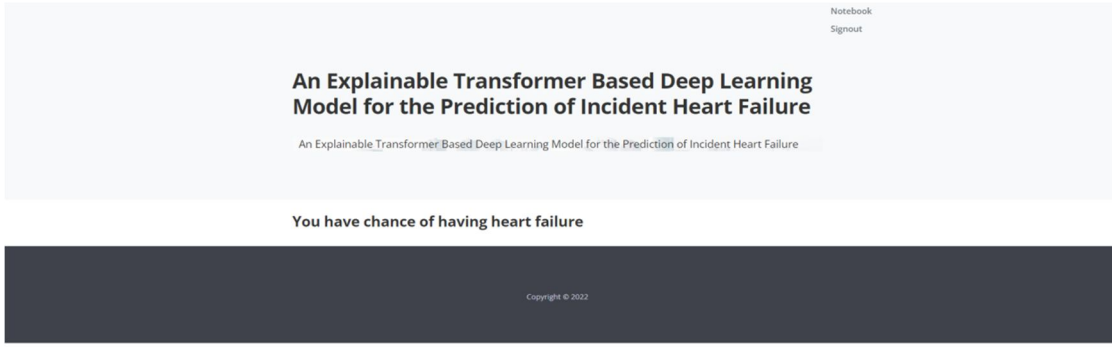


Fig. 8: Output: Chances of Heart Failure

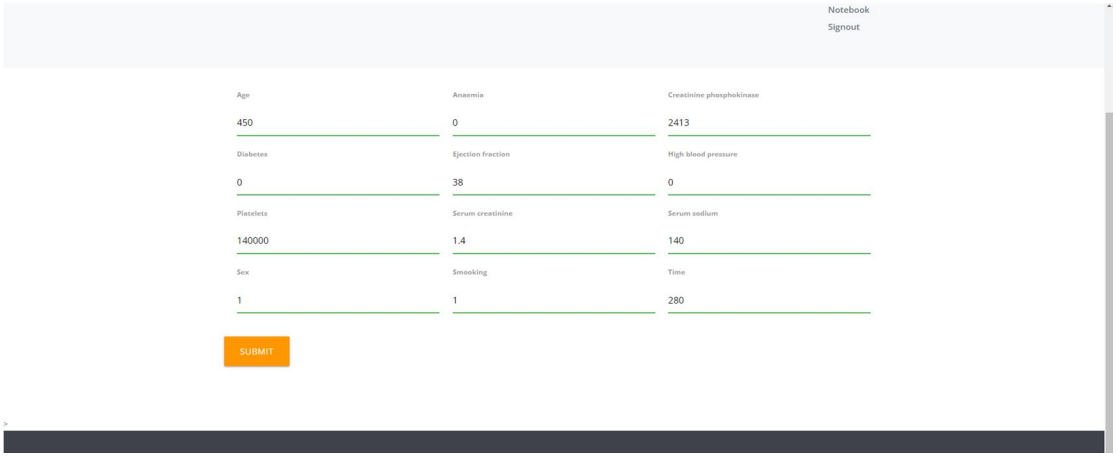


Fig. 9: Another Input for Testing

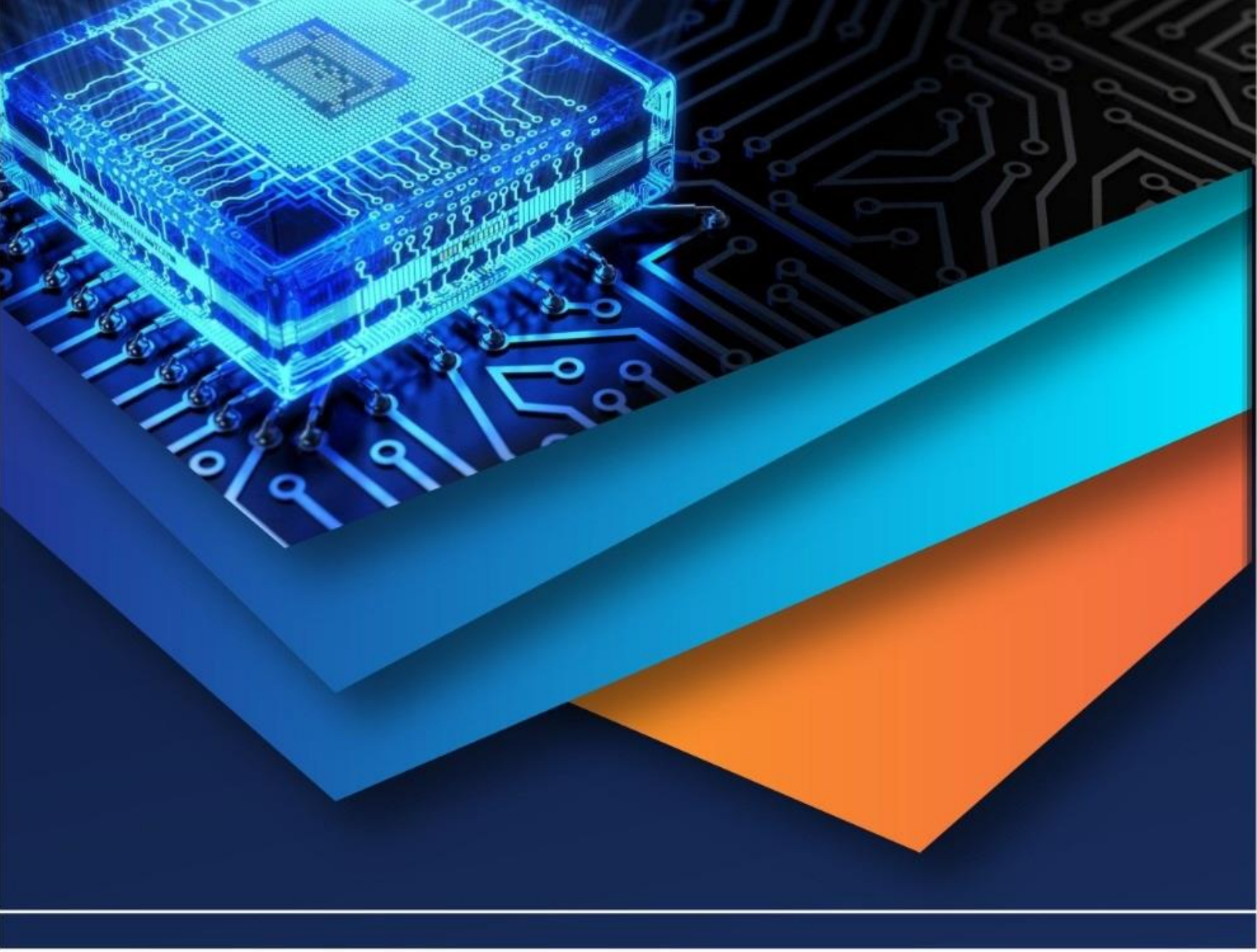
Fig. 10: Output: Safe

VI. CONCLUSION

We developed a superior model for incident HF prediction using routine EHR providing a promising avenue for research into prediction of other complex conditions. Incorporating into routine EHR could alert clinicians to those at risk for more targeted preventive care or recruitment into clinical trials. In addition, we highlight a data-driven approach for identification of potential risk factors that generate new hypotheses requiring causal exploration. We note there are several medications which contribute negatively to HF prediction. Not only are many used to treat established risk factors of HF, but others have not been tested for such an indication and might provide a starting point for drug repurposing studies. The model and analysis could be applied to more deeply phenotype populations for discovery of new disease mechanisms and patterns in other complex conditions.

REFERENCES

[1] B. W. Sahle, A. J. Owen, K. L. Chin, and C. M. Reid, "Risk prediction models for incident heart failure: A systematic review of methodology and model performance," J. Cardiac Failure, vol. 23, no. 9, pp. 680–687, Sep. 2017, doi: 10.1016/j.cardfail.2017.03.005.
[2] E. R. C. Millett and G. Salimi-Khorshidi, "Temporal trends and patterns in mortality after incident heart failure a longitudinal analysis of 86 000 individuals," JAMA Cardiol., vol. 4, pp. 1102–1111, 2019, doi: 10.1001/jamacardio.2019.3593.
[3] N. Conrad et al., "Temporal trends and patterns in heart failure incidence: A population-based study of 4 million individuals," Lancet, vol. 391, no. 10120, pp. 572–580, 2018, doi: 10.1016/S0140-6736(17)32520-5.
[4] F. Rahimian et al., "Predicting the risk of emergency admission with machine learning: Development and validation using linked electronic health records," PLOS Med., vol. 15, no. 11, Nov. 2018, Art. no. e1002695, doi: 10.1371/journal.pmed.1002695.
[5] K. W. Johnson et al., "Artificial intelligence in cardiology," J. Amer. College Cardiol., vol. 71, no. 23, pp. 2668–2679, 2018, doi: 10.1016/j.jacc.2018.03.521.
[6] J. R. A. Solares et al., "Deep learning for electronic health records: A comparative review of multiple deep neural architectures," J. Biomed. Informat., vol. 101, 2020, Art. no. 103337. [Online]. Available: https: //doi.org/10.1016/j.jbi.2019.103337
[7] P. Nguyen, T. Tran, N. Wickramasinghe, and S. Venkatesh, "Deepr: A convolutional net for medical records," IEEE J. Biomed. Health Informat., vol. 21, no. 1, pp. 22–30, Jan. 2017, doi: 10.1109/JBHI.2016.2633963.
[8] E. Choi, M. T. Bahadori, J. A. Kulas, A. Schuetz, W. F. Stewart, and J. Sun, "RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism," in Proc. Int. Conf. Adv. Neural Inf. Process. Syst., 2016, pp. 3512–3520.
[9] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' Explaining the predictions of any classifier," in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2016, vol. 13-17, pp. 1135–1144, doi: 10.1145/2939672.2939778.
[10] D. Smilkov, N. Thorat, B. Kim, F. Viégas, and M. Wattenberg, "SmoothGrad: Removing noise by adding noise," 2017. [Online]. Available: http: //arxiv.org/abs/1706.03825



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)