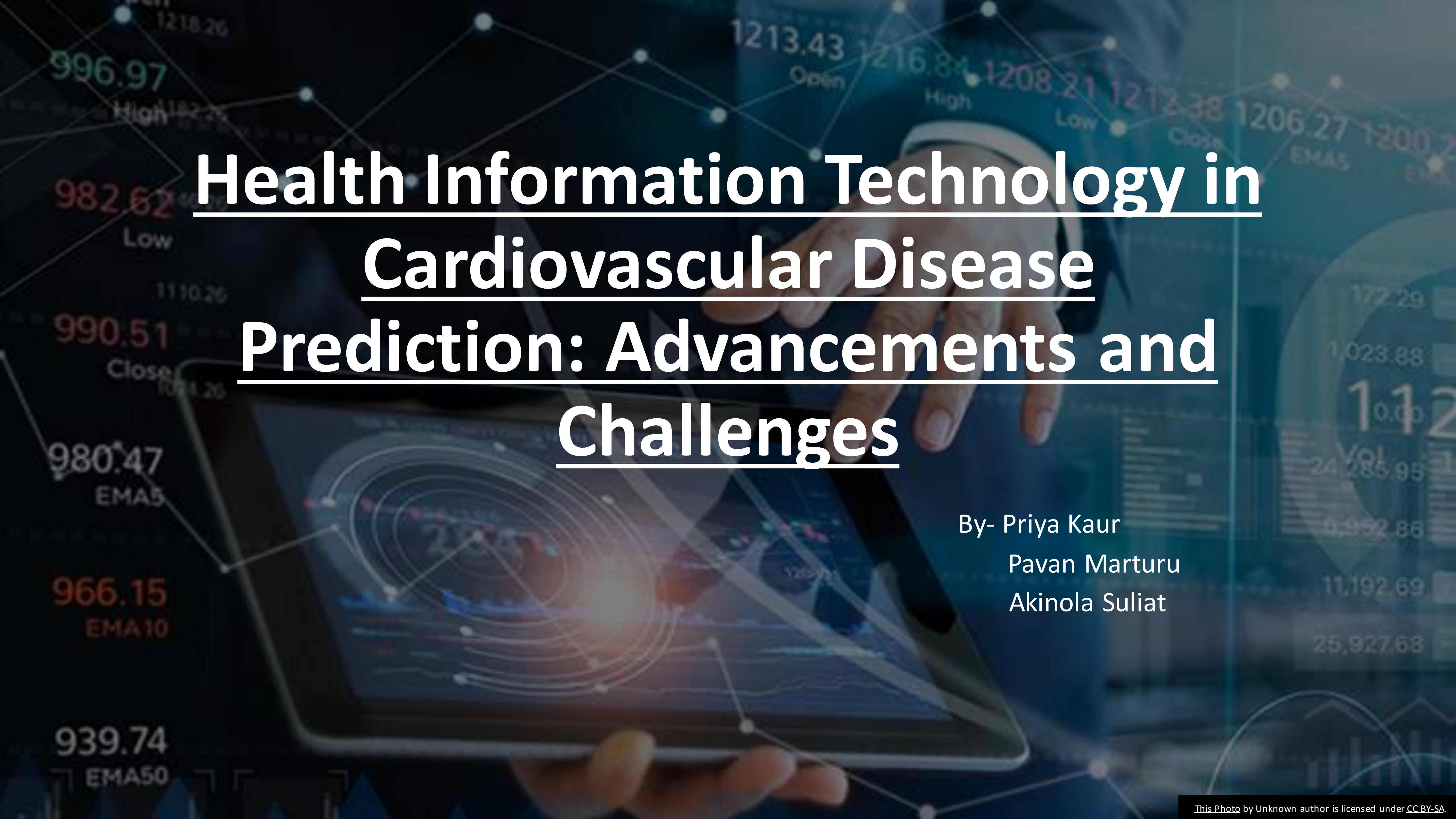


Health Information Technology in Cardiovascular Disease Prediction: Advancements and Challenges

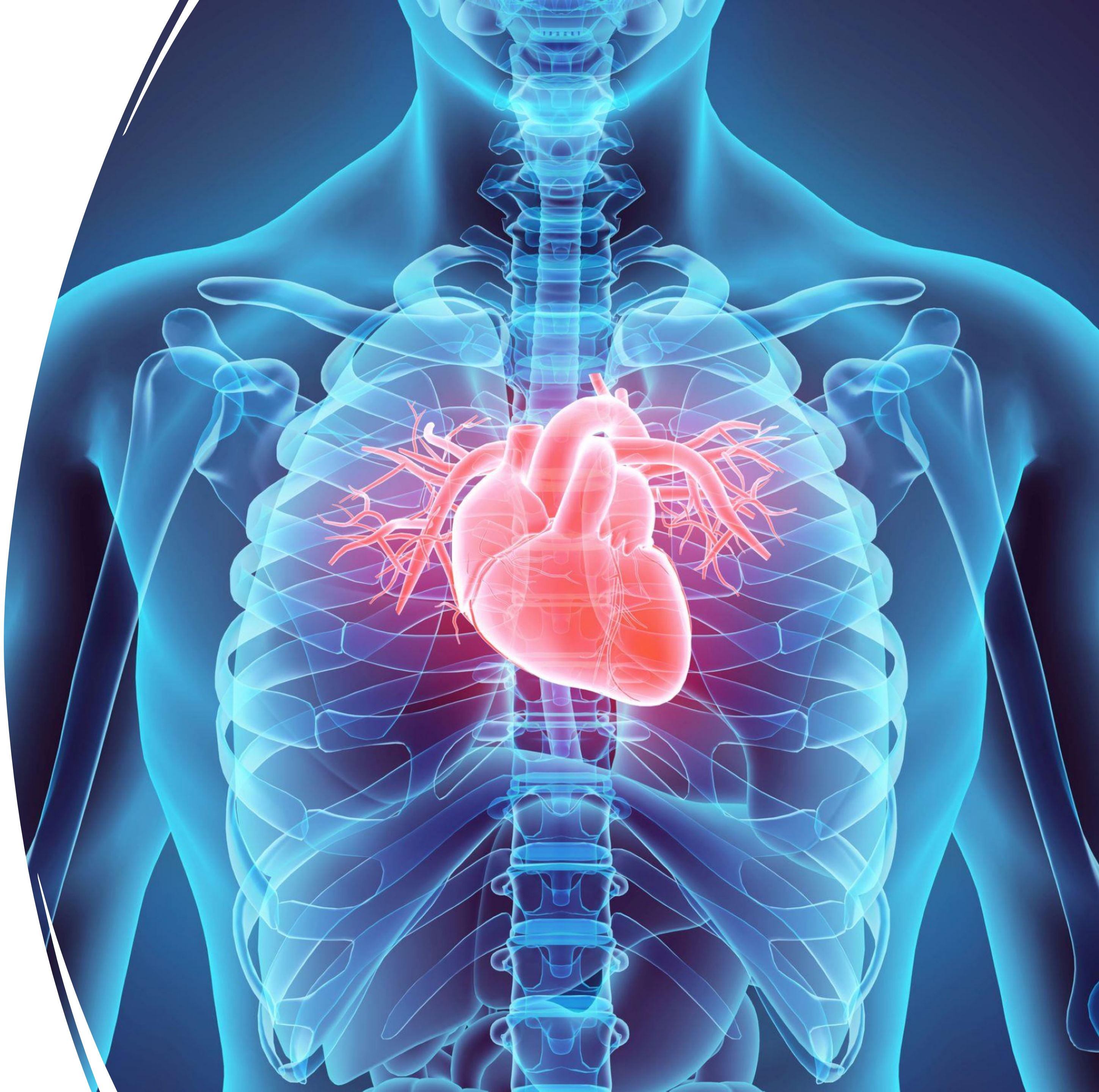
By- Priya Kaur
Pavan Marturu
Akinola Suliat

CONTENTS

- GOAL OF THIS STUDY
- LITERATURE REVIEW
- JUSTIFICATION OF POPULATIONCHOSEN
- ANALYSIS AND INTERPRETATION OF DATA
- LOGIC OF FINDINGS
- REFERENCES

GOAL OF THIS STUDY

- The goal of this study is to look at the promising results of using machine learning to identify patients who are at high risk of developing Cardiovascular disease with or without co-morbidities.
- By identifying and analyzing these patterns, machine learning models can aid in early intervention and hence planning a more personalized treatment plan.



LITERATURE REVIEW

- Population health informatics is an emerging field that utilizes information technology and data analysis techniques to improve health outcomes at the population level
- Machine learning can be used to analyze large datasets of health information and identify patterns, relationships, and insights that can inform the design and implementation of health information systems.
- Machine learning can be used to automate certain aspects of health information management, such as data classification, processing, and analysis, which can improve the accuracy and consistency of health information systems.

STUDIES

Du et al. (2020): Used ML model to predict CVD in hypertension patients using big data and electronic health records with high accuracy and feature selection.

Alaa et al. (2019): Used ML on 400k UK Biobank participants to predict CVD risk, showed potential for large-scale population health architecture, and emphasized feature selection, calibration, and interoperability.

Dinh et al. (2019): Used logistic regression, decision tree, random forest, and gradient boosting for predicting diabetes and CVD with promising results in accuracy and predictive performance.

Goldstein et al. (2017): Used ML for CVD risk prediction with a focus on appropriate feature selection, optimizing model performance, and mitigating issues such as class imbalance and data quality.

Quesada et al. (2019): ML model predicted CVD risk using demographic, clinical, and laboratory data from electronic health records with potential of CNN and RNN and compared to Framingham Risk Score.

Shah et al. (2017): Used ML model to predict heart failure in diabetes patients with high accuracy.

LIMITATIONS

Need for high-quality data, as inconsistencies and missing data can limit effectiveness of ML models.

Need for interpretable models for clinical decision-making.

Potential for bias in ML models from data used and features selected.

Compliance with regulations and guidelines regarding data security and privacy.

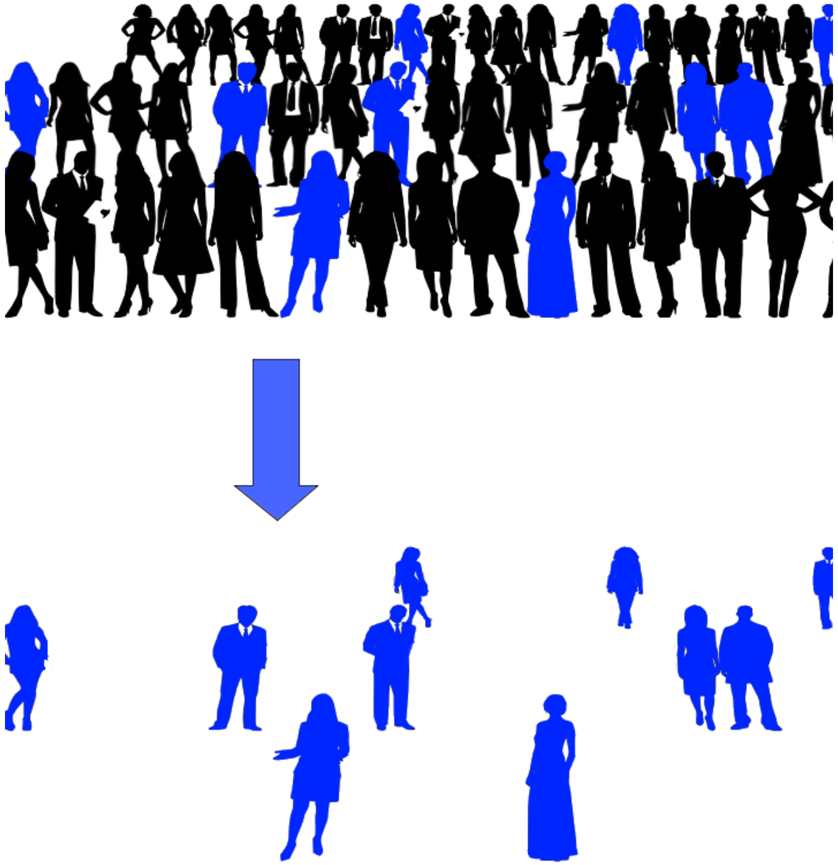
Need for scalable healthcare information architecture to support ML applications

POTENTIAL GAPS

- Limited diversity in study populations and data sources.
- Lack of external validation on independent datasets.
- Limited discussion on clinical implications and integration into current clinical workflow.
- Traditional clinical risk factors considered, with little mention of changes in lifestyle or social determinants of health.

JUSTIFICATION OF POPULATION CHOSEN

- There is no specific age group or demographic that the papers are focusing on.
- All the papers focus on predicting people who are at substantial risk of developing cardiovascular disease.
- Some papers do take other comorbidities such as hypertension into consideration.
- While the papers do not explicitly mention the age group, all the studies do use patients with these conditions who have electronic health records or other clinical data available.

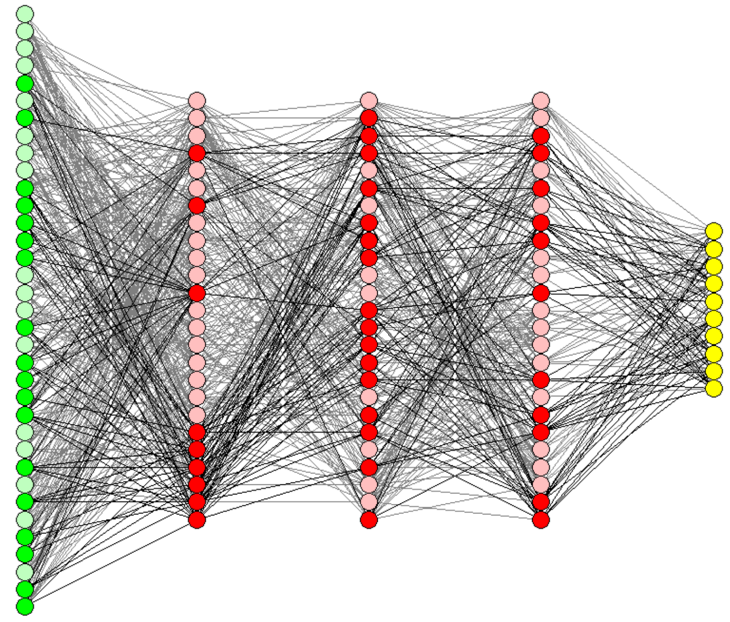


ANALYSIS AND INTERPRETATION OF DATA

- The Literature review aims to provide an overview of various studies that have used machine learning for predicting Cardiovascular disease.
- Goldstein et al. (2017) emphasized the potential of machine learning tools to overcome the limitations of traditional CVD prediction techniques. Quesada et al. (2019) employed machine learning to predict cardiovascular risk based on clinical and laboratory data, whereas Alaa et al. (2019) used data from the UK Biobank to construct an automated machine learning model for CVD risk prediction.
- They also stress the importance of incorporating EHR and big data into machine learning models.

The studies used a range of machine learning models such as decision trees, random forests, support vector machines, artificial neural networks, logistic regression, gradient boosting, and convolutional neural networks.

The results showed that SVM and RF models generally outperformed KNN models in terms of accuracy and AUC. The accuracy rates range from 70% to 80% when trained on large datasets of patient EHRs.



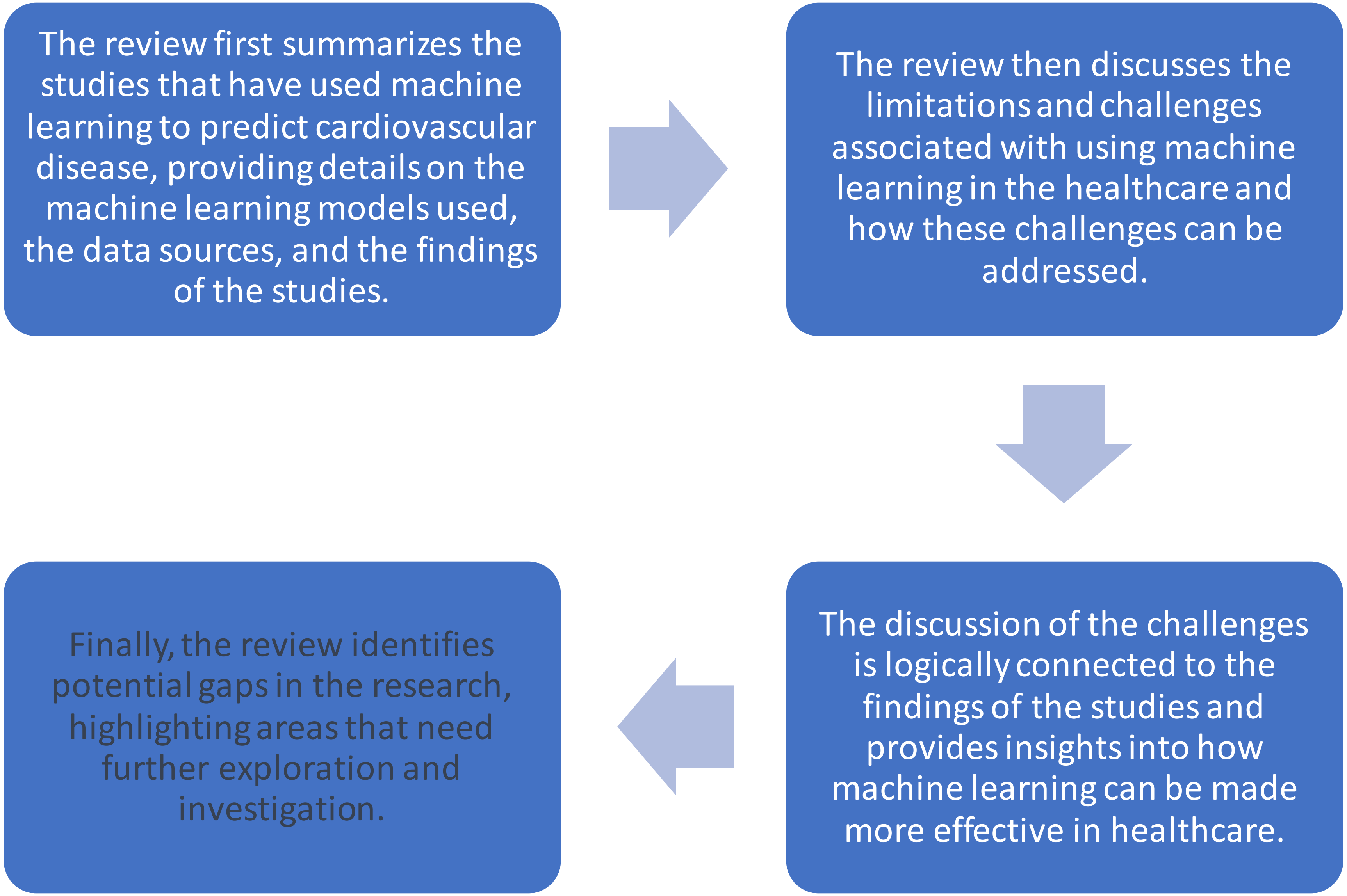
Algorithms	AUC c-statistic	Standard Error*	95% Confidence Interval		Absolute Change from Baseline
			LCL	UCL	
BL: ACC/AHA	0.728	0.002	0.723	0.735	—
ML: Random Forest	0.745	0.003	0.739	0.750	+1.7%
ML: Logistic Regression	0.760	0.003	0.755	0.766	+3.2%
ML: Gradient Boosting Machines	0.761	0.002	0.755	0.766	+3.3%
ML: Neural Networks	0.764	0.002	0.759	0.769	+3.6%

The ACC/AHA risk model served as a baseline for comparison (AUC 0.728, 95% CI 0.723–0.735). All machine-learning algorithms tested achieved statistically significant improvements in discrimination compared to the baseline models (from 1.7% for random forest algorithms to 3.6% for neural networks)

The review also highlights some of the limitations and challenges associated with using machine learning in healthcare:

- high-quality data
- interpretable models
- avoiding biases
- complying with regulations and guidelines regarding data security and privacy
- developing flexible and scalable infrastructure to support machine learning applications.

LOGIC OF FINDINGS


The review first summarizes the studies that have used machine learning to predict cardiovascular disease, providing details on the machine learning models used, the data sources, and the findings of the studies.
The review then discusses the limitations and challenges associated with using machine learning in the healthcare and how these challenges can be addressed.
The discussion of the challenges is logically connected to the findings of the studies and provides insights into how machine learning can be made more effective in healthcare.
Finally, the review identifies potential gaps in the research, highlighting areas that need further exploration and investigation.

REFERENCES

- Du Z, Yang Y, Zheng J, Li Q, Lin D, Li Y, Fan J, Cheng W, Chen XH, Cai Y. Accurate Prediction of Coronary Heart Disease for Patients With Hypertension From Electronic Health Records With Big Data and Machine-Learning Methods: Model Development and Performance Evaluation. JMIR Med Inform. 2020 Jul 6;8(7):e17257. doi: 10.2196/17257. PMID: 32628616; PMCID: PMC7381262.
- Quesada JA, Lopez-Pineda A, Gil-Guillén VF, Durazo-Arvizu R, Orozco-Beltrán D, López-Domenech A, Carratalá-Munuera C. Machine learning to predict cardiovascular risk. Int J Clin Pract. 2019 Oct;73(10):e13389. doi: 10.1111/ijcp.13389. Epub 2019 Aug 4. PMID: 31264310.
- Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J. 2017 Jun 14;38(23):1805-1814. doi: 10.1093/eurheartj/ehw302. PMID: 27436868; PMCID: PMC5837244
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174944 Alaa AM, Bolton T, Di Angelantonio E, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One. 2019 May 15;14(5):e0213653. doi: 10.1371/journal.pone.0213653. PMID: 31091238; PMCID: PMC6519796.
- Dinh, A., Miertschin, S., Young, A. et al. A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. BMC Med Inform Decis Mak 19, 211 (2019). https://doi.org/10.1186/s12911-019-0918-5
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174944
- https://www.jacc.org/doi/abs/10.1016/j.jcmg.2020.08.024

- Eventually, doctors will adopt AI and algorithms as their work partners. This leveling of the medical knowledge landscape will ultimately lead to a new premium: to find and train doctors who have the highest level of emotional intelligence."

  \- **Eric Topol**

THANK YOU