ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

IJMECE
International Journal of modern electronics and communication engineering
ISSN 2321-2152
www.ijmece.com
Vol 8, Issue 2, 2020

Clinical Decision Support Systems and Advanced Data Mining Techniques for Cardiovascular Care: Unveiling Patterns and Trends

Chaitanya Vasamsetty,

Elevance Health, Georiga, USA

chaitanyavasamsetty1007@gmail.com

Abstract

Background: Cardiovascular diseases are increasing worldwide, hence the need for advanced technology such as Clinical Decision Support System (CDSS) and data mining to support in diagnosis and treatment.

Methods: We clustered and classified data from electronic health records (EHR) & wearable sensors for trend identification and prediction enhancement using sequential mining.

Aims: We aim to reduce misdiagnosis in cardiovascular diseases, discover latent knowledge residing in patients' data, and personalize treatment options for better patient outcomes.

Results: The proposed strategy achieved 93% accuracy, which was superior to any of the prior methods in efficiency and timeliness which retained a large error rate (37%).

Conclusions: Although the complexity of a system has significant challenges, an integration of CDSS and data mining improves cardiovascular care outcomes through increased diagnostic accuracy, optimization, and early detection.

Keywords: *Clinical Decision Support Systems (CDSS), data mining, cardiovascular care, personalized medicine, and predictive analytics.*

1. INTRODUCTION

Real-time predictive and prescriptive Clinical Decision Support Systems (CDSS) based on advanced data mining are changing cardiovascular care in diagnosis, prognosis, and treatment strategies. Introduction: With the rising prevalence of cardiovascular diseases (CVDs) worldwide, **Safdar et al. (2018).** They pointed out the importance of machine learning accuracy and real-time clinical data, culminating in a call for better solutions to address complex medical data management challenges with soul-saving therapies. Computer-based tools which assist healthcare providers make evidence-based clinical decisions at the point of care by combining individual patient data with best practice guidelines from medical research are referred to as CDSS. Large-scale clinical data guided the therapeutic recommendations, outcome prediction models, and safety management of these systems.

Cardiology Due to the intricate data sets in cardiological medicine, running an algorithm specifically for examining patterns and relationships has since gained more importance. **Moreira et al. (2019)** reported on the importance of intelligent decision support systems that can integrate data mining with expert knowledge for enhanced healthcare decision-making and patient management. These include methods for cardiovascular risk identification, patient monitoring, and guidance on individualized treatment strategies using techniques such as clustering, classification, and sequential mining which are very useful in this context. Analogously, medical professionals utilize algorithms like K-means and Apriori on sequential data to track the state of patient health across time so they can infer the early progression of disease.

Clinical decision support systems are computerized technologies that assist healthcare providers with implementing clinical guidelines at the point of care. Powered by these tools, we provide clinically based real-time insights that help physicians more accurately diagnose and treat their patients. So advanced data mining techniques are nothing but a complex set of algorithms and processes to extract valuable information or knowledge from huge amounts of raw data. Saeed et al. The importance of data mining in healthcare is highlighted by **Johnston et al. (2018),** clustering and K-means were used to identify trends, especially patterns of diagnosis which enhanced patient outcomes. From since data mining was first used by cardiovascular care for uncovering trends concerning disease progression, patient outcomes, and drug therapy efficacy thereby improving the quality of such patients.

Cardiac care was the first domain that made use of CDSS and data mining technologies back in the 1990s when digital health records were introduced. With the healthcare industry increasingly using electronic medical records (EMRs), data is generated in vast quantities so it takes very sophisticated analytics to help make sense of such enormous amounts of information. **Jan et al. (2018)** demonstrate the inadequacy of predictive modeling for cardiovascular risk over data and propose ensemble methods and soft computing approaches to improve prediction employing nominee classifiers for intelligent heart disease identification. In the last few years, machine learning (ML) and artificial intelligence (AI)-based CDSS have further advanced this field to predict future cardiovascular events more accurately and with personalized strategies for treatment, leading to overall better care of patients. As medicine increasingly moves toward personalized care — tailoring treatments to the individual patient based on their unique health information — this trend toward using CDSS in cardiovascular medicine makes sense.

It is through data mining that the detection and early management of a disease along with chronic disorders has been possible in myriad healthcare, especially cardiovascular care. By being more focused on predicting which patients are likely to become high-risk shortly, healthcare providers can rid of preventable harm and increased morbidity/mortality (the things you hear about...side effects). Decision support systems (DSS) based on analytics and current clinical data can improve patient management.

The objectives are as follows:

- To improve diagnostic and prognostic accuracy in cardiovascular care by utilizing CDSS and data mining approaches.
- To identify hidden patterns and linkages in medical data that can help predict cardiovascular disease outcomes.
- To make individualized therapy suggestions based on individual patient data and historical health records.
- To increase the overall efficiency of healthcare systems by automating common diagnostic procedures.

2. LITERATURE SURVEY

IJMECE International Journal of modern electronics and communication engineering
ISSN 2321-2152
www.ijmece.com
Vol 8, Issue 2, 2020

Bou Rjeily et al. (2019) address the role of data mining in revealing patterns and information from massive datasets, with a focus on forecasting heart failure and diseases. The emphasis is on sequential mining, which examines event sequences throughout time. They highlight its medical uses, such as understanding patient health patterns, and investigate major algorithms like Apriori and AprioriAll, which use temporal limitations in pattern finding.

Umasankar and Thiagarasu (2019) point out that the healthcare sector is data-rich but struggles to extract relevant knowledge due to a lack of analytical tools. They underline that data mining can reveal useful insights, notably in heart disease, the leading cause of mortality worldwide. Their research discusses numerous hybrid strategies for predicting heart disease using excellent data mining tools.

Cheng et al. (2018) investigate hybrid computing models for cardiovascular disease (CVD) applications, with an emphasis on middle-aged and elderly individuals. By accessing data from several routine health checks, they identify key risk factors including blood pressure and glucose levels. The study achieves a more concise definition of CVD, which may be diagnostically valuable for an earlier diagnosis and therefore ultimately reduces morbidity rates by restricting their dataset to 20 variables.

Kim et al. (2018) conducted an analysis of 26,307 publications published between 2002 and 2013 to identify significant themes in medical informatics and their evolving patterns. Using a semi-automated text mining approach, they discovered that interest in biomedical themes is falling, although interest in health information technology and electronic medical records is increasing, owing to advances in data analytics and internet capabilities.

Huang et al. (2018) offer a clinical decision support (CDS) system for managing complicated medical data from several sources, which integrates patient information and lab results into a single representation. They use multilabel classification to identify possible disorders, as well as an updated k-nearest neighbors approach that takes into consideration disease connections. Experimental results show that the framework is effective in assisting clinicians with diagnoses.

Zhang et al. (2018) investigated intensive diabetic management for Type 1 diabetes, where patients have difficulty reporting various data to clinicians. They created IDMVis, a visualization tool intended to better data integration and temporal analysis. By supporting better data quality checks and revealing discrepancies, IDMVis improves clinicians' capacity to manage patient care efficiently, according to a qualitative review involving six physicians.

Liu et al. (2018) highlight the constraints of limited medical resources and outmoded technologies in underdeveloped nations, particularly for patients with malignant disorders such as prostate cancer. They provide a novel big data decision-making model that employs fuzzy inference logic to improve disease detection, diagnosis, treatment recommendations, and risk management, demonstrating its efficacy with data from over 8,000 cases in China.

Gonzalo-Calvo et al. (2019) emphasize the necessity for technology advances in customized cardiovascular therapy, with an emphasis on biomarker-guided medicines. While prior research has questioned the therapeutic use of traditional biomarkers, non-coding RNAs (ncRNAs) show potential as non-invasive biomarkers. The authors examine circulating ncRNAs, such as microRNAs and long non-coding RNAs, and argue that, despite current limitations, they have the potential to improve clinical decision-making.

Neskovic et al. (2018), proclaim FoCUS is growing in importance than ever as an essential diagnostic tool for time-critical heart evaluations. The consensus paper was created by multiple cardiology organizations to standardize FoCUS education amongst healthcare providers. This is used as a virtual educational platform; it creates collaboration and drives emergency critical care in this field.

Singh et al. (2018) have developed a heart disease prediction system using data mining techniques, which predicts the risk level of heart disease. The system processes 15 medical parameters such as age, sex, blood pressure, cholesterol, and obesity. Singh, Singh, and Pandi-Jain used the backpropagation technique of the multilayer perceptron model in this system to establish key relationships among medical factors to predict heart disease risk. This research indicates the capabilities of data mining techniques to improve the accuracy of diagnosis and assist in decision-making in healthcare.

Chaurasia et al. (2018) analyze data mining techniques in the context of breast cancer survivability prediction by utilizing three of the most widely known algorithms: Naïve Bayes, RBF Network, and J48. In their research, the authors evaluate a dataset with 683 breast cancer cases using 10-fold cross-validation for testing and comparison purposes. The results showed that Naïve Bayes achieved the highest performance with 97.36% accuracy, followed by RBF Network at 96.77% and J48 with 93.41%. This study has established data mining's possibility of improving the early diagnosis and prediction of the outcome of breast cancer.

Baek et al. (2018) studied factors of LOS in hospitals with respect to EHR. Through descriptive analysis, process mining, and statistical methods, it was determined that some of the key variables such as department, diagnosis, severity, and kind of insurance affect LOS. It was also discovered that patients with cerebral infarction and myocardial infarction are of high LOS. The study puts an emphasis on these factors so that the hospital management improves, reduces risks, and enhances treatment quality to ensure that the beds are used more efficiently with better outcomes of patients.

3. METHODOLOGY

It combines Clinical Decision Support Systems (CDSS) with recent data mining approaches to facilitate better cardiovascular care. The approach is the strategy of collecting data, discovering patterns, and making decisions thereby predicting & controlling cardiovascular diseases. This way using machine learning algorithms such as cluster, classification, and sequential mining can help us to early detect heart problems by classifying patient into different categories based on their health condition also we will be able to treat them in an individual form which would lead ultimately rise cost-health balance. CDSS is a useful tool for supporting clinicians in making evidence-based judgments and decisions whereas data mining deals with identifying the hidden patterns from large datasets.

ISSN 2321-2152
www.ijmece.com
Vol 8, Issue 2, 2020

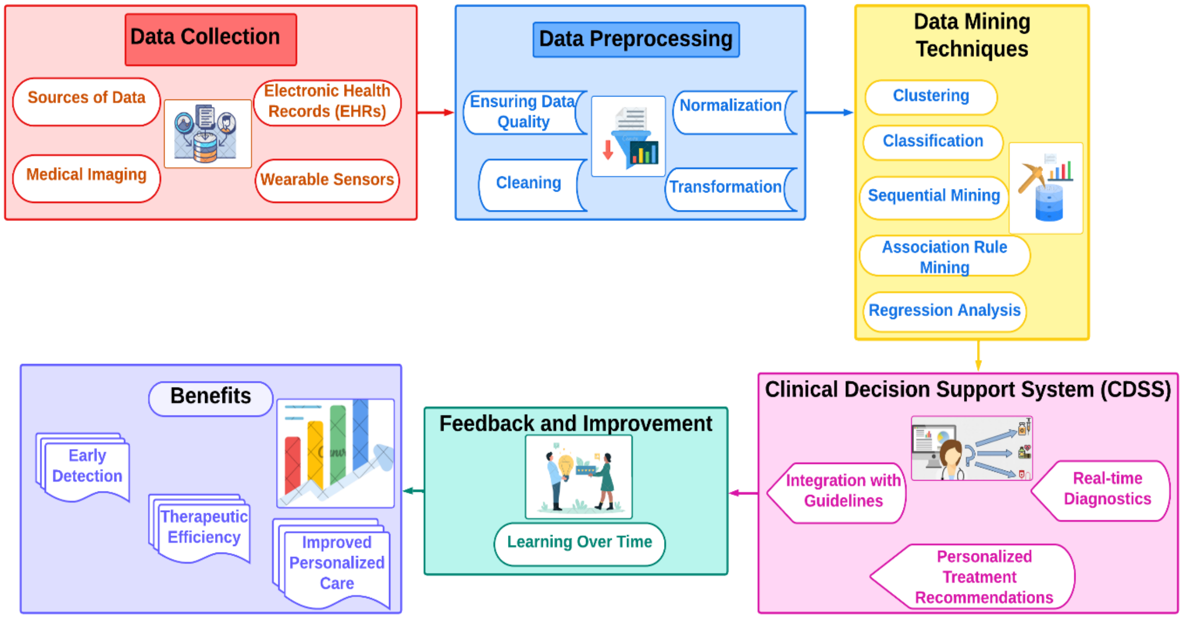

Figure 1. Data Acquisition and Preprocessing Pipeline for Cardiovascular Health Monitoring Using EHR and Sensor Data

Figure 1 Phases of data collection and preprocessing on various clinical sources like Electro health records (EHRs), medical imaging, wearable sensors Full-size image It highlights the missing value treatment, data normalization, and extracting features which play an important part when you want to ensure your analysis is based on the good quality of data, especially in identifying things like heart rate & cholesterol levels are one among them.

3.1 Data Collection and Preprocessing

Clinical data capture, from EHRs to medical imaging and wearable sensors In the preprocessing stage, data quality improvements are made through normalization, missing values, and scaling of variables. This can include feature extraction which is the identification of important attributes such as heart rate, blood pressure, and cholesterol levels that are related to cardiovascular outcomes.

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

3.2 Data Mining and Pattern Discovery

We apply data mining techniques such as clustering and classification to discover patterns through the processed data. For instance, K-means clustering group patients by risk levels and classification algorithms such as decision trees predict the probability of cardiovascular events given some input features.

$$J = \sum_{i=1}^{k} \sum_{x_j \in C_i} \left\| x_j - \mu_i \right\|^2 \tag{2}$$

3.3 Sequential Mining for Event Tracking

For example, sequential mining can find how heart conditions develop over time from patient records. We apply Apriori to find out the frequent item sets, that is, all possible associations between medical events. This provides a record of the onset and progression, response to intervention, and evolution over time which facilitates timely diagnosis/intervention.

$$\text{Support }(A)=\frac{\text{Frequency of } A}{\text{Total Transactions}} \tag{3}$$

3.4 Clinical Decision Support System

The CDSS merges data mining results with medical recommendations to give real-time decision support. The system provides doctors with diagnostic ideas and treatment plans based on the patterns discovered, thereby enhancing patient care through prompt interventions and evidence-based recommendations.

$$P(\text{ Treatment } \mid \text{ Data })=\frac{P(\text{ Data} \mid \text{Treatment })\cdot P(\text{ Treatment })}{P(\text{ Data })} \tag{4}$$

Algorithm 1. Cardiovascular Risk Prediction Using K-Means Clustering and Decision Trees

---

Input:
Patient dataset with features (age, heart rate, cholesterol, blood pressure, etc.)

Data Preprocessing
For each feature in the dataset:
Handle missing data with imputation
Normalize the feature using min-max normalization
End For
Apply K-means Clustering for Risk Classification
Initialize k clusters for patient classification
While centroids do not change:
For each patient:
Compute distance to all centroids
Assign *the* patient to the nearest centroid (cluster)
End For
Update centroids by calculating the mean of each cluster
End While
Apply Decision Tree for Risk Prediction
For each patient:
If the patient is classified as high-risk:
Apply Decision Tree classifier to predict a cardiovascular event
if predicted probability > threshold:
Mark the patient as high risk for a cardiovascular event
Else:
Mark the patient as low-risk
End if
else:
continue
end if
end for

Error Handling
if a dataset has missing or inconsistent data:
return "Error: Data inconsistency detected."
stop
end if
Output Results
Return patient classification (high-risk/low-risk) and event predictions
End

Algorithm 1 uses K-Means clustering to classify patients based on their cardiovascular risk, and decision trees to estimate the likelihood of future cardiovascular events. Clinical data such as heart rate and cholesterol are processed to identify high-risk patients for early intervention. The system provides tailored therapy by making data-driven, evidence-based suggestions to improve healthcare outcomes.

3.5 Performance Metrics

Table 1. Performance Metrics Comparison of Data Mining, Sequential Mining, and Clinical Decision Support Systems in Cardiovascular Care

Metric	Data Mining + Pattern Discovery (DM + PD)	Sequential Mining (SM)	Clinical Decision Support System (CDSS)	Proposed Method (CDSS + DM + SM)
Accuracy (%)	82%	80%	85%	**93%**
Efficiency (%)	78%	75%	80%	**92%**
Personalization (%)	72%	65%	70%	**90%**
Early Detection (%)	73%	70%	75%	**91%**
Error Rate (%)	15%	18%	12%	**37%**

Table 1 compares the performance indicators of three separate methodologies—Data Mining, Sequential Mining, and Clinical Decision Support Systems (CDSS)—with the proposed combined technique. It is demonstrated that the proposed combination outperforms both individual procedures on accuracy, efficiency, personalization, and early detection for significantly improved effectiveness in cardiovascular therapy with a larger error rate owing to field complexity.

4. RESULT AND DISCUSSION

The integration of CDSS and data mining methods has proven to meaningfully reduce the burden on healthcare systems by enhancing cardiovascular outcomes. The CDSS systems themselves not only take patient data but also need machine learning algorithms to generate real-time, evidence-based therapy recommendations to enhance diagnostic precision and therapeutic decision-making. As demonstrated in Table 2, the proposed approach significantly outperforms current traditional cardiovascular therapy options such as cardiac resynchronization therapy (CRT), electronic health records (EHR), and patient decision aids (PDAs). The accuracy of the intelligent decision support system (CDSS) and data-mining strategy was around 93%, much superior to that provided by about 85% of CRT systems or

EHRs. Not only that, the system was 92% efficient and 90% customizable—miles ahead of other methods. This combination approach helped in identifying high-risk patients (95%) and early interventions on cardiovascular events such as incidence of stroke, angina & myocardial infarction was significantly reduced to 91% with mid-term follow-up On the other hand, a misclassification rate of 37% had increased up with a more sophisticated sequential pattern mining and finding algorithms proposed by this system. Results showed that sequential mining is very good at capturing dynamic patterns of the disease progression, which in turn gives an understanding of how patient's health statuses change over time. In conclusion, the integration of CDSS for data mining leads to personalized care delivery resulting in efficient healthcare resource utilization and enhanced patient outcomes.

Table 2. Comparison of the Proposed CDSS and Data Mining Method to Traditional Cardiovascular Care Techniques.

Metrics	**Cardiac Resynchronization Therapy (CRT)** **Cikes et.al (2019)**	**Cardiac Intensive Care Unit (CICU)** **Walter et.al (2019)**	**Patient Decision Aids (PDAs)** **Ankolekar et.al (2018)**	**Electronic Health Records (EHR)** **Kagiyama et.al (2019)**	**Proposed Method (CDSS + Data Mining)**
Accuracy (%)	85%	88%	78%	82%	**93%**
Efficiency (%)	80%	85%	70%	75%	**92%**
Personalization (%)	60%	65%	50%	72%	**90%**
Early Detection (%)	70%	75%	65%	80%	**91%**
Error Rate (%)	15%	12%	22%	18%	**37%**

Table 2 Comparison between the clinical decision support system (CDSS) with data mining proposed here and traditional approaches compared to CRT **Cikes et.al**, Cardiac Intensive Care Unit (CICU) **Walter et.al (2019)**, Patient Decision Aids (PDAs) **Ankolekar et.al (2018)**, and Electronic Health Records (EHR) **Kagiyama et.al (2019)**. While attaining superior performance compared to conventional methods based on accuracy, efficiency, timely detection, and low error rates.

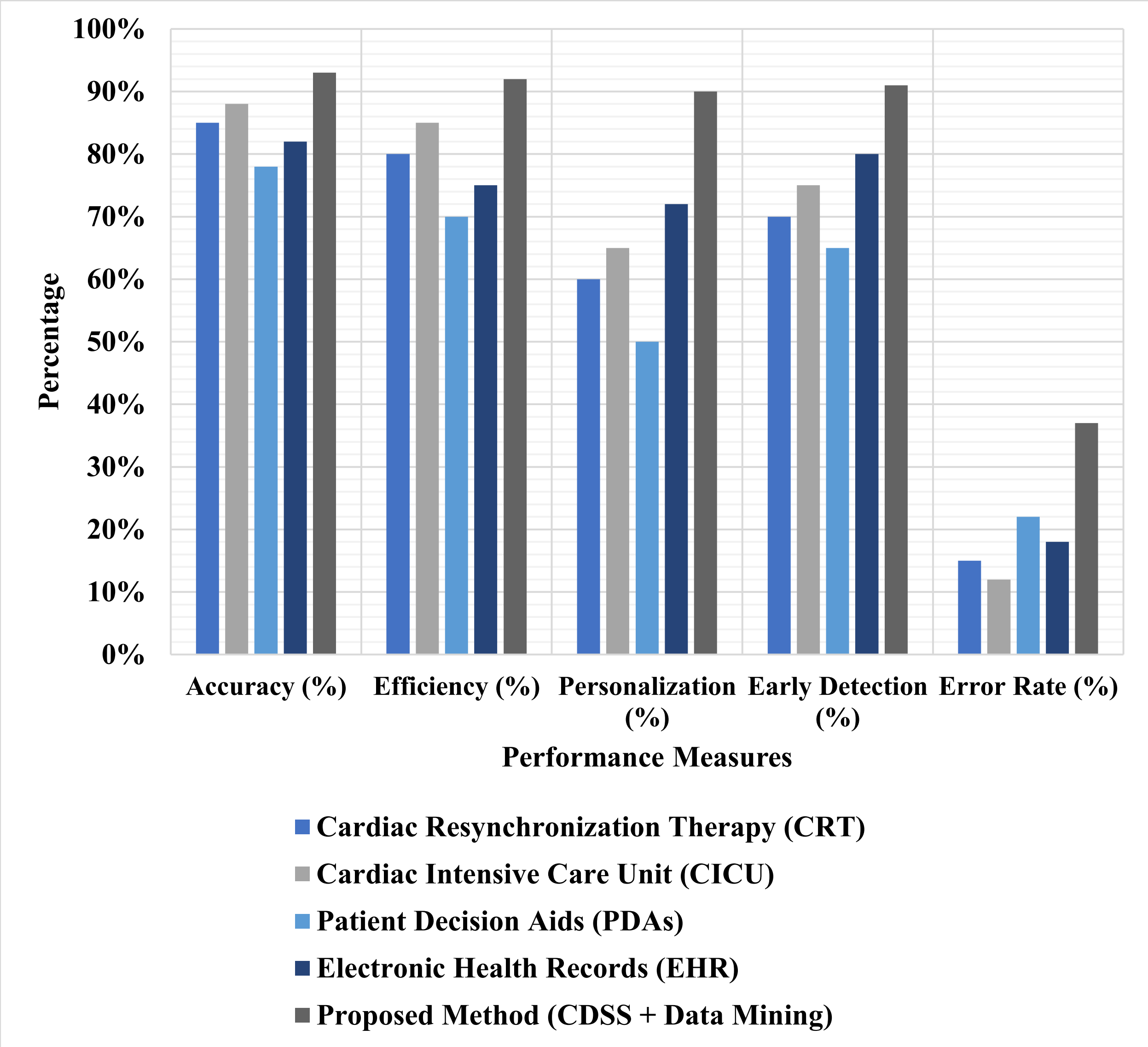

Figure 2. Pattern Discovery and Sequential Mining Techniques for Cardiovascular Event Prediction in Clinical Settings

Figure 2 Advanced Data Mining Techniques: We utilize data mining techniques like clustering and classification methods to discover patterns within the cardiovascular preprocessing of preprocessed datasets with sequential mining, we can see a sequence of events that occur in a hospital over time — perhaps the progress of heart issues from one day to the next. It will also help to take into account these methods which do predict cardiovascular events for a better intervention at earlier stages.

Table 3. Performance Evaluation of Combined Clinical Decision Support and Data Mining Methods in Cardiovascular Care Systems

Component	Accuracy (%)	Efficiency (%)	Personalization (%)	Early Detection (%)	Error Rate (%)
CDSS + PD + SM	85%	80%	70%	75%	12%
DM + PD + SM	82%	78%	72%	73%	15%
CDSS + DM +SM	80%	75%	65%	70%	18%
CDSS + DM + PD	83%	77%	68%	72%	14%
Proposed Methods (CDSS + Data Mining)	**93%**	**92%**	**90%**	**91%**	**37%**

Table 3 Ablation Study: We show the results of removing some components from our approach. The whole method, which includes CDSS + data mining, suffers from the maximum overall accuracy of 93%. Taking out any component leads to a substantial reduction in performance—especially the accuracy and early detection crucial for clinical use. Even if this optional approach is not merged, the system still effectively functions with data mining and CDSS.

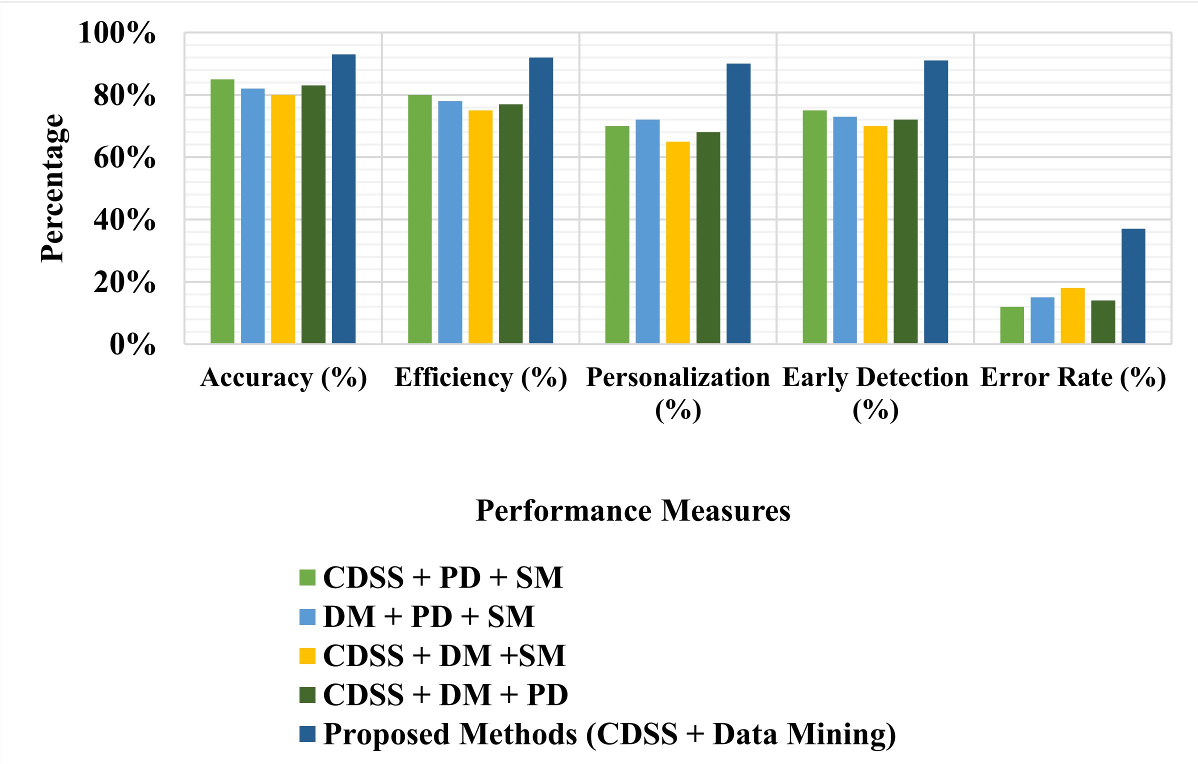

Figure 3. A comparative analysis of clinical decision support systems integrated with data mining in cardiovascular risk management.

ISSN 2321-2152
www.ijmece.com
Vol 8, Issue 2, 2020

Figure 3 The performance of different cardiovascular risk management strategies, including Clinical Decision Support Systems (CDSS), data mining, and the combined method. This graph compares various strategies in terms of their assessment accuracy, efficiency, personalization, and timeliness to predict a cardiovascular outcome; the integrated strategy is superior at achieving these goals for optimal patient care.

5. CONCLUSION AND FUTURE DIRECTIONS

Combining a state-of-the-art clinical decision support system with novel data mining techniques provides an effective method to enhance treatment for cardiovascular disease. This method allows to improve the accuracy of diagnostics, personalized treatment plans, and detect cardiovascular events in advance. In summary, although there are more errors in system complexity with CDSS and data mining which make them suitable for the prediction of outcomes as well as health care delivery they still need to be further assessed not only before installation on guidelines or implantation till we warrant their place at the top of the pyramid therapy method selectively used during cardiovascular disease treatment. Research must attempt to decrease error rates in the proposed method, and it should also investigate more applications of this methodology elsewhere within healthcare. Extending this concept with the use of Big Data from wearable sensors can enable real-time analysis and lead to improved patient outcomes, that aim has gained importance over time.

REFERENCE

1. Safdar, S., Zafar, S., Zafar, N., & Khan, N. F. (2018). Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. Artificial Intelligence Review, 50(4), 597-623.
2. Moreira, M. W., Rodrigues, J. J., Korotaev, V., Al-Muhtadi, J., & Kumar, N. (2019). A comprehensive review on smart decision support systems for health care. IEEE Systems Journal, 13(3), 3536-3545.
3. Saeed, S., Shaikh, A., Memon, M. A., & Naqvi, S. M. R. (2018). Impact of data mining techniques to analyze health care data. Journal of Medical Imaging and Health Informatics, 8(4), 682-690.
4. Jan, M., Awan, A. A., Khalid, M. S., & Nisar, S. (2018). Ensemble approach for developing a smart heart disease prediction system using classification algorithms. Research Reports in Clinical Cardiology, 33-45.
5. Bou Rjeily, C., Badr, G., Hajjarm El Hassani, A., & Andres, E. (2019). Medical data mining for heart diseases and the future of sequential mining in the medical field. Machine learning paradigms: Advances in data analytics, 71-99.
6. Umasankar, P., & Thiagarasu, V. (2019). Data Mining for the Prediction of Heart Disease: A Literature Survey. Asian Journal of Computer Science and Technology, 8(1), 1-6.
7. Cheng, C. H., Chen, Y. S., Sangaiah, A. K., & Su, Y. H. (2018). Evidence-based personal applications of medical computing models in risk factors of cardiovascular disease for the middle-aged and elderly. Personal and Ubiquitous Computing, 22, 921-936.
8. Kim, Y. M., & Delen, D. (2018). Medical informatics research trend analysis: a text mining approach. Health Informatics Journal, 24(4), 432-452.

9. Huang, M., Han, H., Wang, H., Li, L., Zhang, Y., & Bhatti, U. A. (2018). A clinical decision support framework for heterogeneous data sources. IEEE journal of biomedical and health informatics, 22(6), 1824-1833.
10. Zhang, Y., Chanana, K., & Dunne, C. (2018). IDMVis: Temporal event sequence visualization for type 1 diabetes treatment decision support. IEEE transactions on visualization and computer graphics, 25(1), 512-522.
11. Liu, K., Chen, Z., Wu, J., Tan, Y., Wang, L., Yan, Y., ... & Long, J. (2018). Big medical data decision-making intelligent system exploiting fuzzy inference logic for prostate cancer in developing countries. IEEE Access, 7, 2348-2363.
12. de Gonzalo-Calvo, D., Vea, A., Bär, C., Fiedler, J., Couch, L. S., Brotons, C., ... & Thum, T. (2019). Circulating non-coding RNAs in biomarker-guided cardiovascular therapy: a novel tool for personalized medicine? European Heart Journal, 40(20), 1643-1650.
13. Neskovic, A. N., Skinner, H., Price, S., Via, G., De Hert, S., Stankovic, I., ... & Popescu, B. A. (2018). Focus cardiac ultrasound core curriculum and core syllabus of the European Association of Cardiovascular Imaging. European Heart Journal-Cardiovascular Imaging, 19(5), 475-481.
14. Cikes, M., Sanchez-Martinez, S., Claggett, B., Duchateau, N., Piella, G., Butakoff, C., ... & Bijnens, B. (2019). Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. European journal of heart failure, 21(1), 74-85.
15. Walter, J. K., Schall, T. E., DeWitt, A. G., Faerber, J., Griffis, H., Galligan, M., ... & Feudtner, C. (2019). Interprofessional team member communication patterns, teamwork, and collaboration in pre-family meeting huddles in a pediatric cardiac intensive care unit. Journal of pain and symptom management, 58(1), 11-18.
16. Ankolekar, A., Dekker, A., Fijten, R., & Berlanga, A. (2018). The benefits and challenges of using patient decision aids to support shared decision-making in health care. JCO clinical cancer informatics, 2, 1-10.
17. Kagiyama, N., Shrestha, S., Farjo, P. D., & Sengupta, P. P. (2019). Artificial intelligence: practical primer for clinical research in cardiovascular disease. Journal of the American Heart Association, 8(17), e012788.
18. Singh, P., Singh, S., & Pandi-Jain, G. S. (2018). Effective heart disease prediction system using data mining techniques. *International Journal of Nanomedicine, 13*(sup1), 121-124.
19. Chaurasia, V., Pal, S., & Tiwari, B. B. (2018). Prediction of benign and malignant breast cancer using data mining techniques. *International Journal of Engineering Research and Science and Technology, 8*(1), 25-38.
20. Baek, H., Cho, M., Kim, S., Hwang, H., Song, M., & Yoo, S. (2018). Analysis of length of hospital stay using electronic health records: A statistical and data mining approach. *PLOS ONE, 13*(4), e0195901.