**Problem statement and research question**:

An increase in the risk of development of heart diseases due to various reasons is on the surge these days. A sedentary lifestyle and underlying health conditions like diabetes contribute to major reasons associated with this condition. As a preventive measure, people diagnosed with diabetes take prophylactic anti-cardiovascular drugs to prevent complications like atherosclerosis and heart failure. With the advancements in technology, it has become possible to minimize these potential risks through early detection/identification by employing machine learning algorithms on health data. So, with these models, it is possible to identify the potential risk factors contributing to the disease and the vulnerable groups affected by the condition. These models find practical applications in preventing the hospital stay and fatalities associated with these conditions.

With the foregoing in mind, my research question is whether using machine learning algorithms on EHR data improves the prediction of cardiovascular disease in the general population with diabetes as compared to conventional periodic examinations.

**Literature Review:**

Citation of the article:

Segar, M. W., Vaduganathan, M., Patel, K. V., McGuire, D. K., Butler, J., Fonarow, G. C., Basit, M., Kannan, V., Grodin, J. L., Everett, B., Willett, D., Berry, J., & Pandey, A. (2019). Machine Learning to Predict the Risk of Incident Heart Failure Hospitalization Among Patients With Diabetes: The WATCH-DM Risk Score. *Diabetes care*, *42*(12), 2298–2306. <https://doi-org.proxy.ulib.uits.iu.edu/10.2337/dc19-0587>

The article I reviewed in the literature addresses the same question I wish to study. Here a random survival forest model was developed using 8,756 patients’ health data to identify the predictors for the development of incident Heart Failure. Later the model was validated for efficiency by testing on a cohort of individuals suffering from type-II diabetes. In addition, the performance was compared to the existing best performing cox-based model.

The model employed in this study is able to outperform the existing model in terms of identifying the risks and scoring them based on severity.

**Method:**

I'd want to create and build a machine learning model called "CVS Risk" to possibly determine the risk of cardiovascular disease in a diabetic population and compare its efficacy to an existing method. To do this, I'd want to develop cohort experiment that allows for longer-term follow-up of the individuals. Because the algorithm predicts the occurrence based on identifying characteristics, my independent variable is the model, and my dependent variable is cardiovascular disease.

The algorithm, termed "CVS Risk," was created using logistic regression and was trained using EHR data from 10,000 diabetic patients. The decision trees classifier and random forest classifier are two more algorithms that were put to the test. However, as compared to the "CVS Risk," these models perform poorly on the training dataset. With this training, the developed model is able to identify the potential risk factors contributing to the incidence of cardiovascular disease. With these factors, a risk-scoring algorithm is deployed to classify patients into different groups like high risk, low risk, and average risk groups.

To externally validate this model, a cohort of 15,000 people aged 55 to 75 years old was randomly selected and followed up on for five years. Sensitivity, specificity, positive predictive value, and negative predictive value are all used to evaluate the model's effectiveness. The accuracy is compared to a previously published model. Significant improvements were seen in the ROC curves and confidence interval values.

This method of detecting risk variables allows for prognostic treatments to reduce the likelihood of hospitalization and death. Both the patient and the doctor benefit from this new approach.

The suggested approach is also useful for clinical decision support.

Heuristic Evaluation

Name: familydoctor.org

Link / Web address: <https://familydoctor.org/>

Based on my observation of the site, every display begins with an appropriate heading and stylistic concerns. The icon designs are evident throughout but there is no visual feedback on the website. When the cursor is placed on the text of interest, statements of errors and identification pop up. The language in which the icons are written is more of simple terms that can be easy to interpret by the user. When it comes to the menu, each topic is unique and details to answer the user’s questions. The sequential steps are logically connected to prompt the user with the meaningful information they need. The user has the freedom to return to the previous state as the site is flexible. The symbols and letters are clearly defined. Each window is separately labeled with appropriate titles.

The overall display of the site is quite simple. The color contrast is accurate to the facility's bright display. There is tricolor representation with gentle sounds for the video presentations if any. There are multiple windows present and switching between them is easy. The only limitation of this site is the lack of a feedback mechanism for the user if they have performed any error.

There are various available options to browse information regarding the topic of interest to the user. All the information is provided in (A-Z) order making it easy for the user to browse. Additionally, the webpage provides accurate contact information. The whole content of the site is given in two languages- English and Spanish. The privacy policy is in accordance with the American Academy of Family Physicians (AAFP) Privacy Policy.Graphical user interface, website

Description automatically generated

Graphical user interface, website

Description automatically generatedA picture containing text, screenshot, person

Description automatically generatedGraphical user interface, website

Description automatically generatedGraphical user interface, website

Description automatically generatedGraphical user interface, website

Description automatically generatedGraphical user interface, application

Description automatically generatedGraphical user interface, text, website

Description automatically generatedGraphical user interface, website

Description automatically generatedGraphical user interface, application, website

Description automatically generated