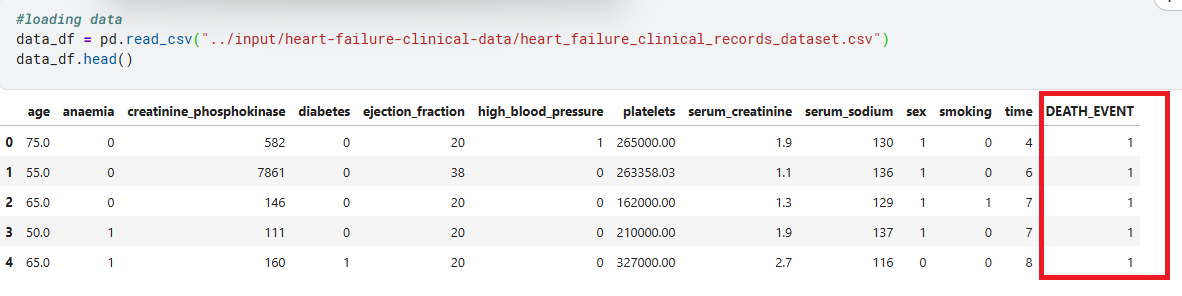
# Section 1: Description

Purpose: Create & compare different ML (Machine Learning) algorithms for generating models for predicting Heart Failure event. This dataset contains the characteristics of patients with & without heart failure events.

Figure 1 Heart Failure Dataset

## Goals: This dataset will be used to train & test models and algorithms used to predict probability of Death Event or Heart Failure event. The Label for Dataset is highlighted in Red as shown in Figure 1 Heart Failure Dataset. The Label indicates if the heart failure is predicted to occur 1 (Potential Death event) or 0 (Survival or no adverse event for a negative outcome). Early diagnosis is key to surviving a heart failure event & ML gives the ability to analyze complex patterns.

# Section 2: Responsibilities

Requirements – Harish Lakshman & Shalini k Nistala. Describes what concerns ML algorithm addresses.

Design – Steven Gonzales. How would we solve this problem. Design a manual solution to tackle the problem.

Implementation - All Team members contributing to Software, ML algorithm by collaborating on Kaggle.

Testing - All Team members contributing to Software, ML algorithm by collaborating on Kaggle.

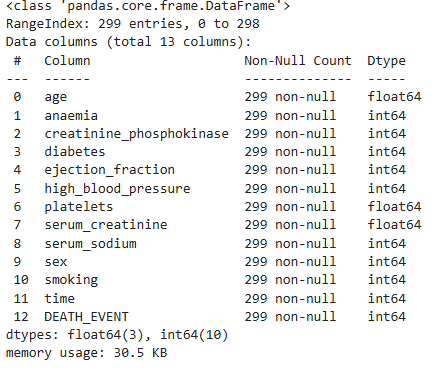
Report - All Team members contributing to Report.

Presentation - Harish Lakshman. Will prepare presentation.

Project Status – Harish Lakshman

# Section 3: Information about the Data (https://www.kaggle.com/code/harishlakshman/heart-failure-prediction-svm-and-ann/edit)

Figure 2 Statistics for Dataset for Heart failure detection

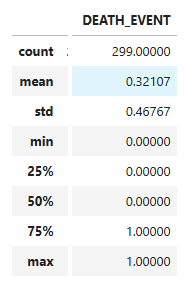
A screenshot of a computer

AI-generated content may be incorrect.

Figure 3 Distribution of Dataset

A group of colorful text

AI-generated content may be incorrect.



# Section 4: Project Status

Progress:

1. Identified requirements
2. Downloaded & pre-processed data set.
3. Currently analyzing data set in preparation for modelling.

Methods & Planned tasks:

1. Finish data exploration
2. Try PCA on the data
3. Try KNN clustering
4. Try linear & logistic regression
5. Attempt decision tree
6. Implement mild imbalance, explore SMOTE, weighted loss function or threshold tuning.
7. Find hyper-parameters using Grid-search, random search or Bayesian optimization.
8. Compare SVM(Support vector machines) versus ANN(Artificial neural networks) across different metrics(AUC-ROC, precision-recall, F1-score).
9. Try ensemble approach (e.g.: SVM + ANN + Boosting method).
10. Advantages/Disadvantages of new techniques in comparison with previously learned methods.
11. Attempt Convolutional Neural Networks & Deep Learning.