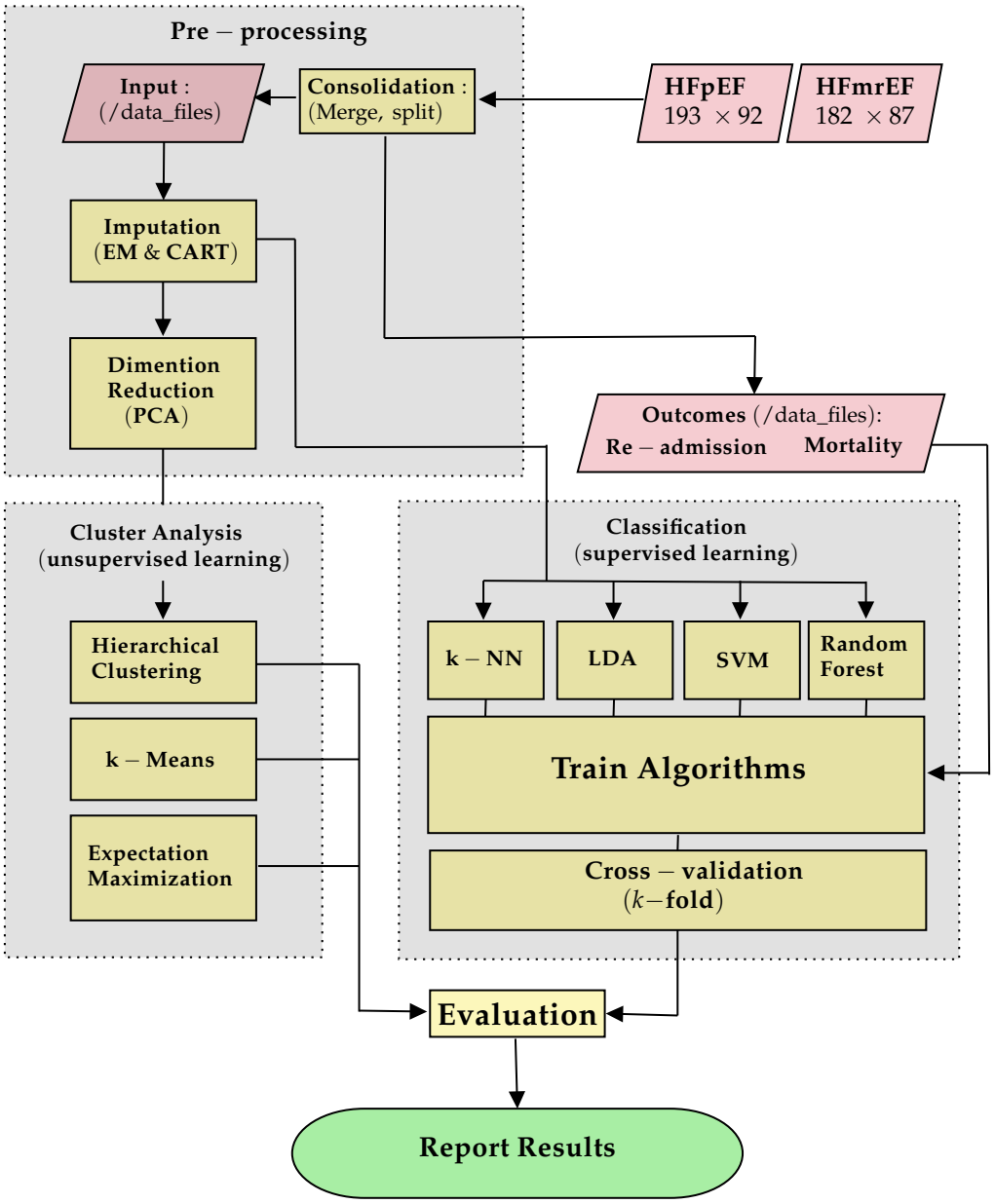
Overview and Plan for Deadline

Samir Adrik

October 2018

1 Research Overview



2 Clustering Analysis

Research question:

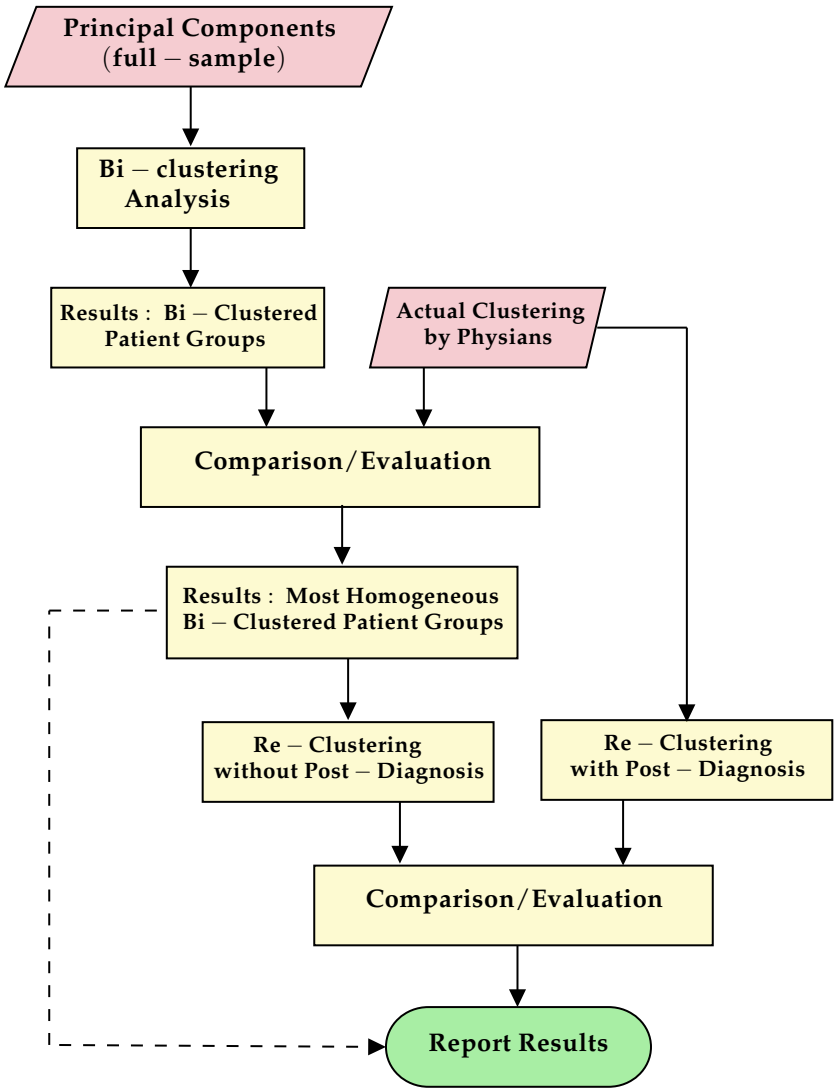
How well do various clustering methods (Hierarchical clustering, K-Means and Expectation–Maximization) perform in producing phenotypically distinct clinical patient groups (i.e. phenomapping) with HFpEF and HFmrEF?

Design:

We first look at the Bi-clustering problem where we rank algorithms based on how unique the patient groups they produce are compared to the diagnosis done by the physicians. We assume here that the patient groups can only be HFmrEF or HFpEF, i.e. only two clusters. The evaluation criteria is the number of statistically significant baseline characteristics. This is calculated using using the Person χ^2 test for categorical variables, ANOVA for normally distributed variables and Kruskal–Wallis test for non-normally distributed variables.

Next, we look at how well the various clustering algorithms do in producing new clustering groups within the subtypes of HF (both where the diagnosis was done by the physicians and the 'best' performing algorithm from the Bi-clustering problem mentioned previously).

The following process flow shows the design of the experiments in this section:



Results:

Tabell 1: Number of significant baseline characteristics

$C = 2$	Bi-clustering, i.e. two clusters			
Actual			59	
Hierarchical			62	
K-Means			62	
EM			54	
$C = 3$	With Post-Diagnosis		Without Post-Diagnosis	
	HFpEF	HFmrEF	HFpEF	HFmrEF
Hierarchical	55	53	48	51
K-Means	49	53	48	53
EM	56	44	42	42

Conclusion:

We have reason to believe that the ML algorithms Hierarchical and K-Means clustering can be very useful in producing patient groups that are more phenotypically unique given that the objective is to challenge the diagnosis of the physicians. Similar results are reported in the literature (see chapter 2 in thesis). We have commented on the baseline characteristics for each patient groups generated by the clustering analysis (see chapter 4).

However, if the objective is to use the results to find additional ”new clusters”, we cannot say with certainty that the choice of clustering algorithms or the clustering data used, i.e. weather its with or without post-diagnosis will systematically enhance the "uniqueness"of the patient groups. The results vary with the imputation method used, the number of principal components and the sample size.

3 Classification

Research question:

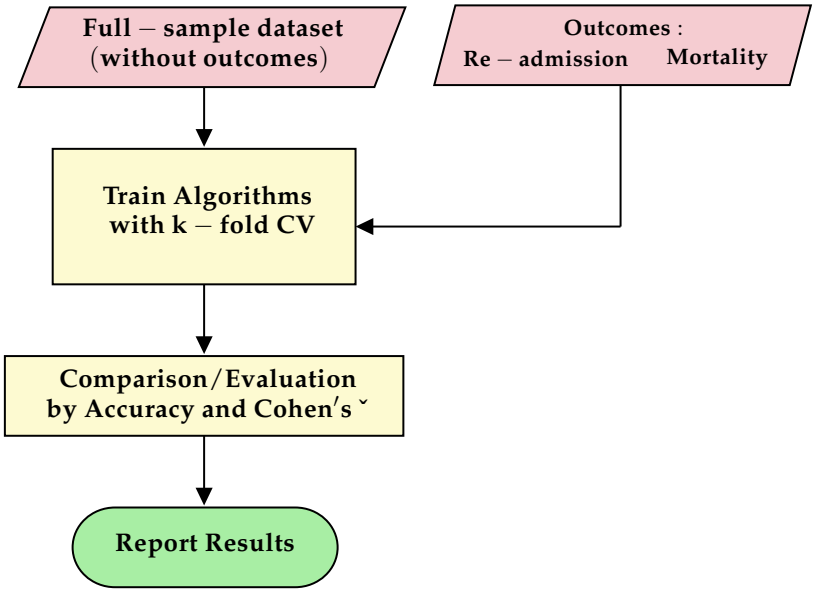
How well can various classification algorithms (K-nearest neighbours, Linear Discriminant Analysis, Support Vector Machines and Random Forest) predict the clinical outcomes (mortality and re-admission) of patients with HFpEF and HFmrEF?

Design:

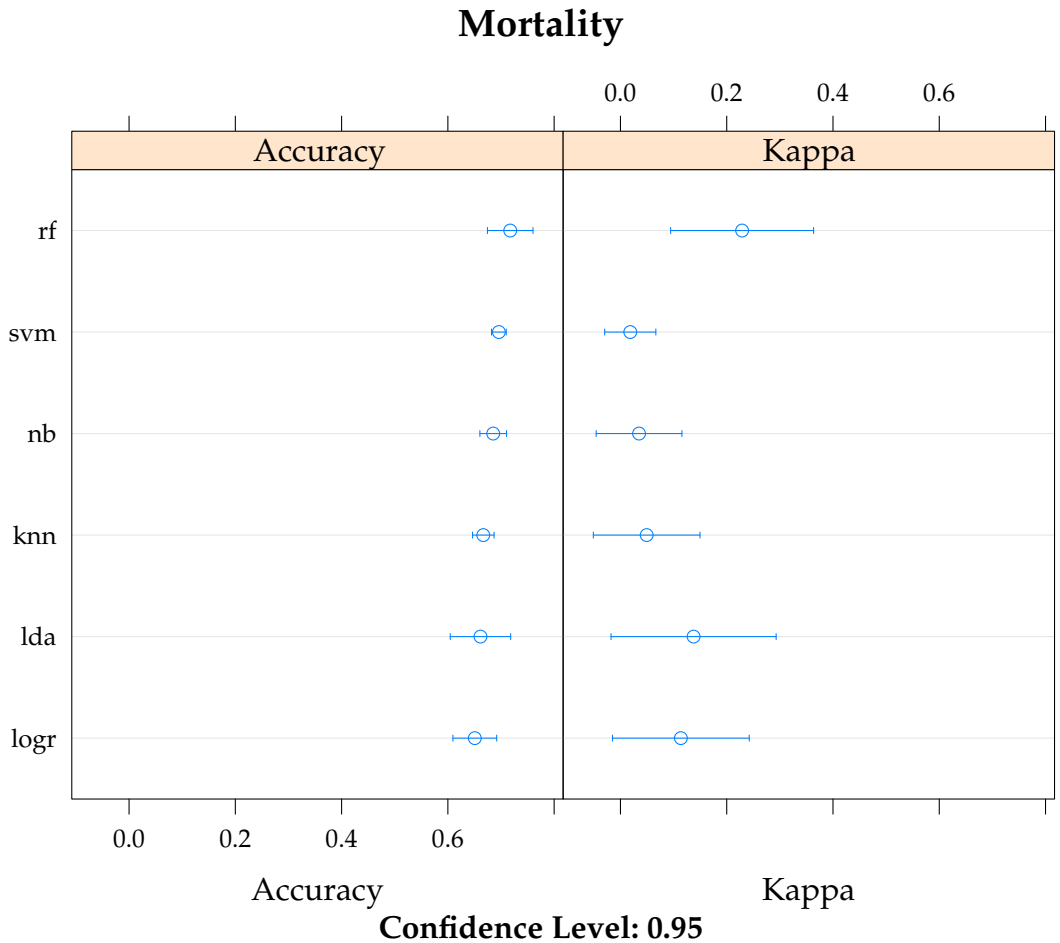
Here we run the imputed full sample dataset through the classification algorithms and train them using k-fold cross validation ($k = 5$). The evaluation criteria is that of the accuracy, i.e. the proportion of true results and the Cohen’s kappa defined by:

$$\kappa \equiv \frac{p_0 - p_e}{1 - p_e} \tag{1}$$

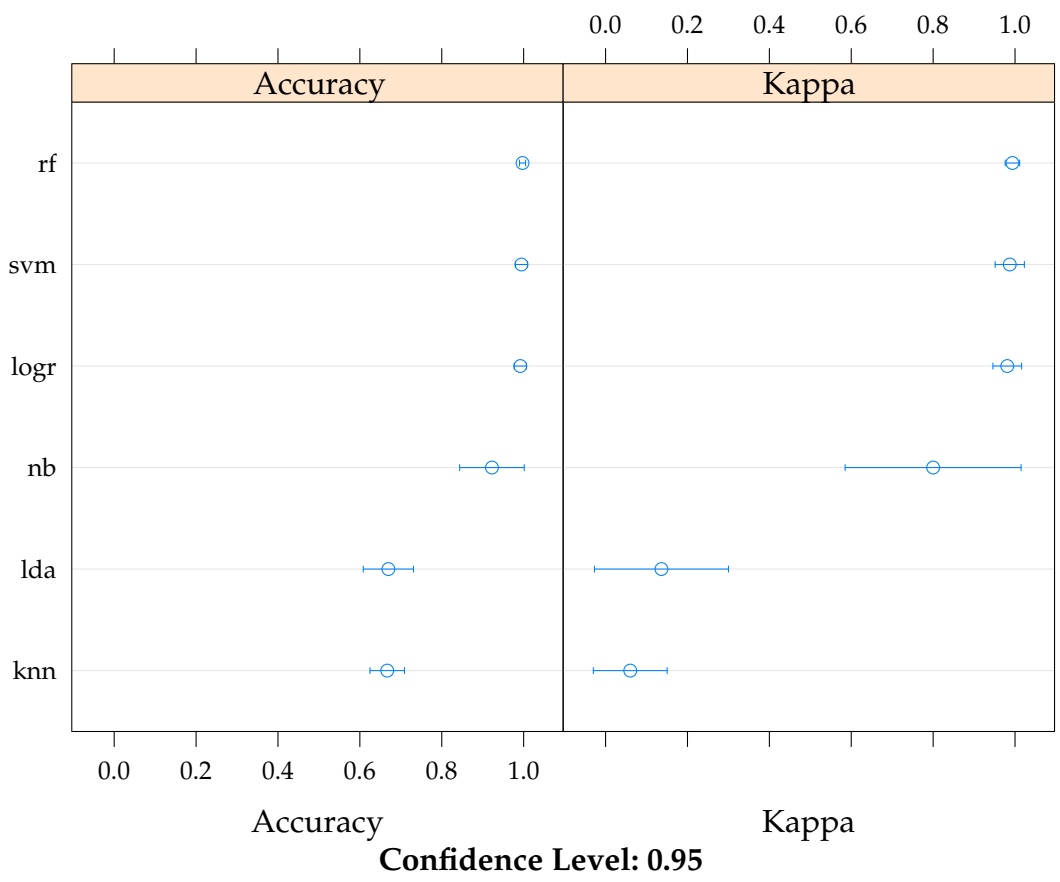
where p_0 is the accuracy given by $ACC = (TP + TN)/(P + N)$, and $p_e = 1/N^2 \sum_k n_{k1} n_{k2}$, where k is the number of categories / classes, N the number of items and n_{k1} the number of times rater i predicted category k. The statistical learning problem is that of a Binary classification problem given by weather re-admission / mortality occurred (`TRUE`) or not (`FALSE`). The following process flow illustrated the design of the experiments in this section.



Results:



Confidence Level: 0.95



Conclusion:

We have reasons to believe that classification algorithms can with varying degree of success predict both the mortality rate and the re-admission of HF patients. The overall most accurate classification algorithm is that of the Random Forest (71% for mortality and 98% for re-admission). This algorithm also presents the highest values for the Cohen κ. There is large variation in the classification results as can be seen in the two plots above.

We need to emphasize that the classification results are very sensitive to the number of folds used in the cross-validation. However, the results presented are in-line with much of the literature presented in chapter 2 of the thesis.

4 Delivery Plan

- Wrap-up chapter 4 - Deals with the experiments, finished all the coding. Just need to comment on the results related to the classification in more detail. **Expected Deadline:** 1. November.
- Finish conclusive remarks, abstract / summary and further expand the appendix with supplementary reading / results. **Expected Deadline:** 10. November.
- Proofreading **Expected Deadline:** 20. November.
- Hand-in thesis **Expected Deadline:** 25. November.

5 Action Points

- Add current version of thesis to `bitbucket repository` in pdf so that Ulf and Kristin can see / review the thesis as the final sections are completed.
- Give status updates on the deadlines presented above to both Ulf and Kristin as they occur.
- Schedule final meeting(s) before delivery.