Levels of 1,2,3,4 could imply a step relationship, change to 5 1:0 flag variables to reduce the imply so computer doesn’t think 4 is 4 time smore than 1

Not just mean, need SD, could split the time variable in smaller chunks (window the data)

Imputation, what to do with imputation, mean value imply average patient thing are missing at random ( no it is more likely to be missing systematically)

maybe find neigest neighbour,

Encode the missing ness, add missing ness variable and impute for it

We scaled all paremeters, we have values between 0 to 1,

Variable in different scales need to be adjusted so everything sort of have same weight in terms of numbers

We need to minimising Euclidean distance, so each value able is normalised for the computation, other wie variables with bigger values will influence more

Higher AUC, better sensitivity/specitivity ratio