Disease Prediction

PRANJAL SRIVASTAVA

2022-03-30

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.1.2
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.1.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(magrittr)
```

```
## Warning: package 'magrittr' was built under R version 4.1.2
```

```
##
## Attaching package: 'magrittr'
```

```
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.1.2
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.1.2
```

```
## Loading required package: lattice
```

```
library(Boruta)
```

```
## Warning: package 'Boruta' was built under R version 4.1.2
```

```
library(rpart)
```

```
## Warning: package 'rpart' was built under R version 4.1.2
```

```
library(rpart.plot)
```

```
## Warning: package 'rpart.plot' was built under R version 4.1.2
```

```
library(pROC)
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

D. Data And Experiment

Data Preparation and Cleaning

The most important part of this project is to import and clean the data as needed.The dataset contains the variables as various clinical symptoms and prognosis as a result of combination of symptoms. The data is originally taken from Kaggle data source: [https://www.kaggle.com/datasets/kaushil268/disease-prediction-using-machine-learni
(https://www.kaggle.com/datasets/kaushil268/disease-prediction-using-machine-learning)

Importing data

We set the working directory as we have already downloaded the 'Disease.csv' data in my folder from the website.

```
setwd("/Users/pranjalsrivastava/Desktop/Projects/Disease_Prediction_Model")
```

After setting the working directory, we imported the `csv` data file and generating the raw data frame "Disease1"

```
Disease1 <- read.csv('Disease.csv')

Disease1 <- Disease1 %>% select(-X)


Disease1[] <- lapply(Disease1, as.factor)
```

We can clearly see that there no missing values in our final dataframe. Prognosis has 42 unique categorical values names as various diseases. All other variables are valued either 1 or 0

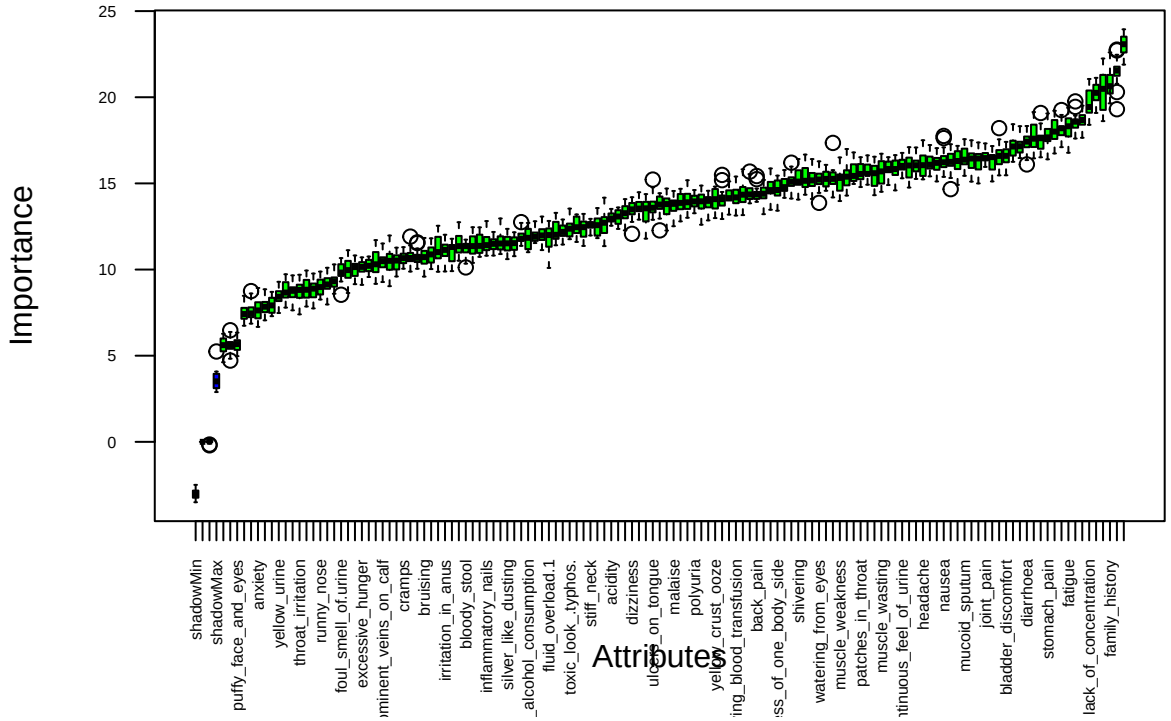
Feature Selection (Boruta)

```
Boruta <- Boruta(prognosis~., data = Disease1, doTrace = 2)
```

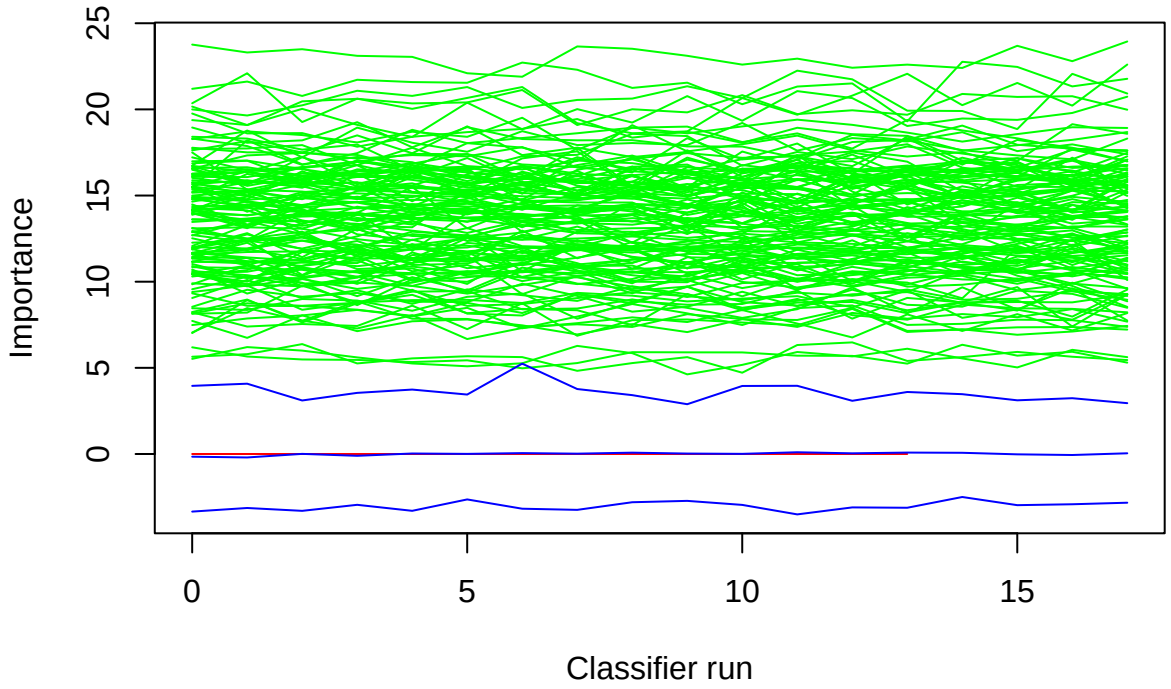
```
##  1. run of importance source...

##  2. run of importance source...

##  3. run of importance source...

##  4. run of importance source...

##  5. run of importance source...

##  6. run of importance source...

##  7. run of importance source...

##  8. run of importance source...
```

```
##  9. run of importance source...

##  10. run of importance source...

##  11. run of importance source...

##  12. run of importance source...

##  13. run of importance source...

##  14. run of importance source...

## After 14 iterations, +60 secs:

##  confirmed 130 attributes: abdominal_pain, abnormal_menstruation, acidity, acute_liver_failure, alte

##  rejected 1 attribute: fluid_overload;

##  still have 1 attribute left.

##  15. run of importance source...

##  16. run of importance source...

##  17. run of importance source...

##  18. run of importance source...

## After 18 iterations, +1.4 mins:

##  confirmed 1 attribute: puffy_face_and_eyes;

##  no more attributes left.
```

```
plot(Boruta, las = 2, cex.axis = 0.5)
```




```
plotImpHistory(Boruta)
```




```
attStats(Boruta)
```

```
##                                meanImp medianImp    minImp    maxImp
## itching                      20.205122 20.255305 19.088752 21.125202
## skin_rash                    14.687857 14.612328 13.471400 15.867553
```

```
## nodal_skin_eruptions           18.046187 18.007509 16.737949 19.212937
## continuous_sneezing            13.623133 13.538891 12.898583 14.486071
## shivering                      15.289066 15.107457 14.061299 16.413457
## chills                         14.870945 14.784199 14.162112 15.773883
## joint_pain                     16.516806 16.470960 15.958053 17.320315
## stomach_pain                   17.792447 17.671431 16.396868 19.058945
## acidity                        13.015820 12.960748 12.544515 13.518716
## ulcers_on_tongue               13.694797 13.584306 12.889317 15.221930
## muscle_wasting                 15.669884 15.696239 14.513739 16.721540
## vomiting                       20.399151 20.487740 18.614565 22.248255
## burning_micturition            15.181647 15.175284 14.099351 16.411376
## spotting_.urination            17.685016 17.595413 16.453652 19.139541
## fatigue                        18.207508 18.314792 16.779014 19.387773
## weight_gain                     5.599838  5.612910  4.621594  6.274480
## anxiety                         7.643983  7.626452  6.672173  8.930135
## cold_hands_and_feets            5.622558  5.622808  4.717834  6.476577
## mood_swings                     7.474511  7.417686  6.741075  8.476038
## weight_loss                    15.096487 15.094446 14.119256 16.201103
## restlessness                    9.127263  9.136143  7.973190 10.102890
## lethargy                       10.240826 10.214530  9.567898 10.756153
## patches_in_throat              15.634125 15.544447 14.787463 16.517632
## irregular_sugar_level           7.940304  7.920313  7.293936  8.703909
## cough                          14.054976 14.087236 12.932815 15.494457
## high_fever                     18.106121 18.200258 17.103848 19.255985
## sunken_eyes                    16.236823 16.249102 14.666995 17.272204
## breathlessness                 16.317504 16.444527 15.103378 17.500590
## sweating                       16.313065 16.300652 14.913816 17.178280
## dehydration                    15.845092 15.834670 14.594233 16.640620
## indigestion                    11.496122 11.533146 10.775938 12.161114
## headache                       16.145158 16.057377 15.236091 17.504047
## yellowish_skin                 13.948319 13.905084 13.006408 15.182143
## dark_urine                     17.189538 17.157915 16.469907 18.319764
## nausea                         16.327940 16.225515 15.378590 17.754205
## loss_of_appetite               14.492596 14.361077 13.844740 15.683865
## pain_behind_the_eyes           16.138872 16.081900 15.410335 17.227718
## back_pain                      14.417281 14.361880 13.931657 15.427047
## constipation                   12.227932 12.183266 11.452068 12.903030
## abdominal_pain                 15.279370 15.262140 14.572451 16.075340
## diarrhoea                      17.315956 17.425682 16.100275 18.329986
## mild_fever                     18.676055 18.678714 17.647450 19.514111
## yellow_urine                    8.454498  8.421938  7.476941  9.289412
## yellowing_of_eyes              16.573832 16.596089 15.466303 18.206840
## acute_liver_failure             9.963223  9.967905  8.639921 11.122202
## fluid_overload                  0.000000  0.000000  0.000000  0.000000
## swelling_of_stomach            11.499691 11.373280 10.737522 12.448024
## swelled_lymph_nodes            11.410447 11.362209 10.386636 12.455620
## malaise                        13.721903 13.813762 12.515341 14.561331
## blurred_and_distorted_vision    9.332043  9.319725  8.596628 10.296197
## phlegm                         11.913484 11.789089 11.375928 12.750886
## throat_irritation               8.761289  8.817528  7.399713  9.698219
## redness_of_eyes                 8.821902  8.893422  7.745286  9.769752
## sinus_pressure                  8.715955  8.665477  7.793139  9.726287
## runny_nose                      9.004320  8.983501  8.260576  9.881359
## congestion                      8.884158  8.828043  7.854155  9.941998
```

```
## chest_pain                       17.184605 17.140626 16.253962 18.435625
## weakness_in_limbs                13.035959 13.078596 12.358488 13.602600
## fast_heart_rate                  14.414107 14.401584 13.213824 15.201792
## pain_during_bowel_movements      10.900027 10.879197 10.027363 11.811599
## pain_in_anal_region              11.032559 11.309525  9.907723 11.785625
## bloody_stool                     11.163199 11.360943 10.123528 11.734532
## irritation_in_anus               11.064031 11.141835  9.889266 12.019273
## neck_pain                        15.713004 15.433796 14.967011 16.796518
## dizziness                        13.443720 13.492474 12.071526 14.196190
## cramps                           10.655583 10.663356 10.048511 11.300151
## bruising                         10.693411 10.720675  9.592241 11.833389
## obesity                           9.986171 10.143199  9.151278 10.816632
## swollen_legs                     10.724181 10.681036 10.046512 11.576241
## swollen_blood_vessels             7.867412  7.844359  7.052510  8.854811
## puffy_face_and_eyes               5.636393  5.675927  4.973212  6.339953
## enlarged_thyroid                 10.430402 10.504996  9.035734 11.975813
## brittle_nails                    10.493181 10.523431  9.596620 11.269826
## swollen_extremeties              10.425132 10.306166  9.171227 11.707818
## excessive_hunger                 10.091864 10.153678  9.110888 10.706669
## extra_marital_contacts           15.307479 15.263276 14.187765 17.349163
## drying_and_tingling_lips          7.607925  7.452505  6.863179  8.749858
## slurred_speech                   13.362931 13.553875 11.794823 14.377605
## knee_pain                        12.625743 12.716100 11.359525 14.171124
## hip_joint_pain                   12.603086 12.577032 11.824837 13.567874
## muscle_weakness                  15.186200 15.315523 13.974821 16.505238
## stiff_neck                       12.580109 12.566958 11.977757 12.941711
## swelling_joints                  12.070423 11.978734 11.129363 13.075532
## movement_stiffness               13.245900 13.271972 12.274580 14.087172
## spinning_movements               12.513484 12.458010 11.602743 13.199110
## loss_of_balance                  16.280082 16.524004 15.115403 17.167485
## unsteadiness                     17.530084 17.620246 16.514808 19.085475
## weakness_of_one_body_side        14.636659 14.616553 13.416115 15.451447
## loss_of_smell                     8.700061  8.766726  7.633280  9.661264
## bladder_discomfort               16.564354 16.598509 15.472464 17.409955
## foul_smell_of.urine               9.873646  9.826023  8.537531 10.666400
## continuous_feel_of_urine         16.051859 16.010428 15.296768 16.770063
## passage_of_gases                 15.443631 15.389359 14.294165 16.997260
## internal_itching                 15.475607 15.591822 14.315492 16.547014
## toxic_look_.typhos.              12.384441 12.370139 11.762782 13.537984
## depression                       10.726836 10.669993  9.883358 11.911043
## irritability                     12.247945 12.016205 11.409135 13.289802
## muscle_pain                      23.020866 23.077932 21.894102 23.940344
## altered_sensorium                21.501727 21.547001 19.305500 22.762548
## red_spots_over_body              15.260059 15.149664 14.343851 16.673148
## belly_pain                       12.360365 12.459704 11.370678 13.230159
## abnormal_menstruation            13.840982 13.986719 12.854193 14.570824
## dischromic._patches              18.541540 18.582844 17.604486 19.759630
## watering_from_eyes               15.241204 15.225155 13.871586 16.279664
## increased_appetite               13.675523 13.800079 12.263150 14.700928
## polyuria                         13.967749 13.954439 13.268863 14.614368
## family_history                   20.747934 20.642540 19.646492 22.602127
## mucoid_sputum                    16.485717 16.375826 15.441880 17.565464
## rusty_sputum                     16.326319 16.446016 14.998539 17.375789
## lack_of_concentration            19.708693 19.432235 18.398341 21.061195
```

```
## visual_disturbances            11.085682 11.014722  9.872130 12.540892
## receiving_blood_transfusion    14.266945 14.187624 13.135380 15.398582
## receiving_unsterile_injections 14.331192 14.343640 13.159805 15.817755
## coma                           15.951794 16.039159 14.760000 17.101353
## stomach_bleeding               16.146746 16.191078 15.115360 17.045835
## distention_of_abdomen          11.622757 11.511123 10.899158 12.965698
## history_of_alcohol_consumption 11.810475 11.842004 11.007505 12.706576
## fluid_overload.1               11.806383 12.002379 10.094805 12.820788
## blood_in_sputum                13.903652 13.960802 12.602411 14.873257
## prominent_veins_on_calf        10.474254 10.488302  9.274715 11.511526
## palpitations                   13.736151 13.770773 12.273593 14.430074
## painful_walking                11.976898 11.905232 11.537846 12.802369
## pus_filled_pimples             15.868446 15.834287 14.903849 16.520320
## blackheads                     16.030117 16.053630 15.259784 17.083785
## scurring                       15.694316 15.548758 14.924326 16.626781
## skin_peeling                   11.488694 11.492882 10.811451 12.181334
## silver_like_dusting            11.505301 11.520791 10.604971 12.361083
## small_dents_in_nails           11.477019 11.360780 10.514161 12.747409
## inflammatory_nails             11.532988 11.417046 10.804546 12.812734
## blister                        14.262035 14.145707 13.190027 15.113263
## red_sore_around_nose           13.869620 13.868637 12.796019 14.685415
## yellow_crust_ooze              14.046284 14.085010 12.986261 14.814348
##                                 normHits  decision
## itching                        1.0000000 Confirmed
## skin_rash                      1.0000000 Confirmed
## nodal_skin_eruptions           1.0000000 Confirmed
## continuous_sneezing            1.0000000 Confirmed
## shivering                      1.0000000 Confirmed
## chills                         1.0000000 Confirmed
## joint_pain                     1.0000000 Confirmed
## stomach_pain                   1.0000000 Confirmed
## acidity                        1.0000000 Confirmed
## ulcers_on_tongue               1.0000000 Confirmed
## muscle_wasting                 1.0000000 Confirmed
## vomiting                       1.0000000 Confirmed
## burning_micturition            1.0000000 Confirmed
## spotting_.urination            1.0000000 Confirmed
## fatigue                        1.0000000 Confirmed
## weight_gain                    1.0000000 Confirmed
## anxiety                        1.0000000 Confirmed
## cold_hands_and_feets           1.0000000 Confirmed
## mood_swings                    1.0000000 Confirmed
## weight_loss                    1.0000000 Confirmed
## restlessness                   1.0000000 Confirmed
## lethargy                       1.0000000 Confirmed
## patches_in_throat              1.0000000 Confirmed
## irregular_sugar_level          1.0000000 Confirmed
## cough                          1.0000000 Confirmed
## high_fever                     1.0000000 Confirmed
## sunken_eyes                    1.0000000 Confirmed
## breathlessness                 1.0000000 Confirmed
## sweating                       1.0000000 Confirmed
## dehydration                    1.0000000 Confirmed
## indigestion                    1.0000000 Confirmed
```

```
## headache                          1.0000000 Confirmed
## yellowish_skin                    1.0000000 Confirmed
## dark_urine                        1.0000000 Confirmed
## nausea                            1.0000000 Confirmed
## loss_of_appetite                  1.0000000 Confirmed
## pain_behind_the_eyes              1.0000000 Confirmed
## back_pain                         1.0000000 Confirmed
## constipation                      1.0000000 Confirmed
## abdominal_pain                    1.0000000 Confirmed
## diarrhoea                         1.0000000 Confirmed
## mild_fever                        1.0000000 Confirmed
## yellow_urine                      1.0000000 Confirmed
## yellowing_of_eyes                 1.0000000 Confirmed
## acute_liver_failure               1.0000000 Confirmed
## fluid_overload                    0.0000000  Rejected
## swelling_of_stomach               1.0000000 Confirmed
## swelled_lymph_nodes               1.0000000 Confirmed
## malaise                           1.0000000 Confirmed
## blurred_and_distorted_vision      1.0000000 Confirmed
## phlegm                            1.0000000 Confirmed
## throat_irritation                 1.0000000 Confirmed
## redness_of_eyes                   1.0000000 Confirmed
## sinus_pressure                    1.0000000 Confirmed
## runny_nose                        1.0000000 Confirmed
## congestion                        1.0000000 Confirmed
## chest_pain                        1.0000000 Confirmed
## weakness_in_limbs                 1.0000000 Confirmed
## fast_heart_rate                   1.0000000 Confirmed
## pain_during_bowel_movements       1.0000000 Confirmed
## pain_in_anal_region               1.0000000 Confirmed
## bloody_stool                      1.0000000 Confirmed
## irritation_in_anus                1.0000000 Confirmed
## neck_pain                         1.0000000 Confirmed
## dizziness                         1.0000000 Confirmed
## cramps                            1.0000000 Confirmed
## bruising                          1.0000000 Confirmed
## obesity                           1.0000000 Confirmed
## swollen_legs                      1.0000000 Confirmed
## swollen_blood_vessels             1.0000000 Confirmed
## puffy_face_and_eyes               0.9444444 Confirmed
## enlarged_thyroid                  1.0000000 Confirmed
## brittle_nails                     1.0000000 Confirmed
## swollen_extremeties               1.0000000 Confirmed
## excessive_hunger                  1.0000000 Confirmed
## extra_marital_contacts            1.0000000 Confirmed
## drying_and_tingling_lips          1.0000000 Confirmed
## slurred_speech                    1.0000000 Confirmed
## knee_pain                         1.0000000 Confirmed
## hip_joint_pain                    1.0000000 Confirmed
## muscle_weakness                   1.0000000 Confirmed
## stiff_neck                        1.0000000 Confirmed
## swelling_joints                   1.0000000 Confirmed
## movement_stiffness                1.0000000 Confirmed
## spinning_movements                1.0000000 Confirmed
```

```
## loss_of_balance                1.0000000 Confirmed
## unsteadiness                   1.0000000 Confirmed
## weakness_of_one_body_side      1.0000000 Confirmed
## loss_of_smell                  1.0000000 Confirmed
## bladder_discomfort             1.0000000 Confirmed
## foul_smell_of.urine            1.0000000 Confirmed
## continuous_feel_of_urine       1.0000000 Confirmed
## passage_of_gases               1.0000000 Confirmed
## internal_itching               1.0000000 Confirmed
## toxic_look_.typhos.            1.0000000 Confirmed
## depression                     1.0000000 Confirmed
## irritability                   1.0000000 Confirmed
## muscle_pain                    1.0000000 Confirmed
## altered_sensorium              1.0000000 Confirmed
## red_spots_over_body            1.0000000 Confirmed
## belly_pain                     1.0000000 Confirmed
## abnormal_menstruation          1.0000000 Confirmed
## dischromic._patches            1.0000000 Confirmed
## watering_from_eyes             1.0000000 Confirmed
## increased_appetite             1.0000000 Confirmed
## polyuria                       1.0000000 Confirmed
## family_history                 1.0000000 Confirmed
## mucoid_sputum                  1.0000000 Confirmed
## rusty_sputum                   1.0000000 Confirmed
## lack_of_concentration          1.0000000 Confirmed
## visual_disturbances            1.0000000 Confirmed
## receiving_blood_transfusion    1.0000000 Confirmed
## receiving_unsterile_injections 1.0000000 Confirmed
## coma                           1.0000000 Confirmed
## stomach_bleeding               1.0000000 Confirmed
## distention_of_abdomen          1.0000000 Confirmed
## history_of_alcohol_consumption 1.0000000 Confirmed
## fluid_overload.1               1.0000000 Confirmed
## blood_in_sputum                1.0000000 Confirmed
## prominent_veins_on_calf        1.0000000 Confirmed
## palpitations                   1.0000000 Confirmed
## painful_walking                1.0000000 Confirmed
## pus_filled_pimples             1.0000000 Confirmed
## blackheads                     1.0000000 Confirmed
## scurring                       1.0000000 Confirmed
## skin_peeling                   1.0000000 Confirmed
## silver_like_dusting            1.0000000 Confirmed
## small_dents_in_nails           1.0000000 Confirmed
## inflammatory_nails             1.0000000 Confirmed
## blister                        1.0000000 Confirmed
## red_sore_around_nose           1.0000000 Confirmed
## yellow_crust_ooze              1.0000000 Confirmed
```

```
Boruta
```

```
## Boruta performed 18 iterations in 1.352149 mins.
##  131 attributes confirmed important: abdominal_pain,
## abnormal_menstruation, acidity, acute_liver_failure, altered_sensorium
## and 126 more;
```

```
##  1 attributes confirmed unimportant: fluid_overload;
```

Removing "fluid_overload" as it has been rejected in feature selection

```
Disease <- Disease1 %>% select(-fluid_overload)
```

Unique Diseases(prognosis) in the dataset.

```
unique(Disease$prognosis)
```

```
##  [1] Fungal infection
##  [2] Allergy
##  [3] GERD
##  [4] Chronic cholestasis
##  [5] Drug Reaction
##  [6] Peptic ulcer diseae
##  [7] AIDS
##  [8] Diabetes
##  [9] Gastroenteritis
## [10] Bronchial Asthma
## [11] Hypertension
## [12] Migraine
## [13] Cervical spondylosis
## [14] Paralysis (brain hemorrhage)
## [15] Jaundice
## [16] Malaria
## [17] Chicken pox
## [18] Dengue
## [19] Typhoid
## [20] hepatitis A
## [21] Hepatitis B
## [22] Hepatitis C
## [23] Hepatitis D
## [24] Hepatitis E
## [25] Alcoholic hepatitis
## [26] Tuberculosis
## [27] Common Cold
## [28] Pneumonia
## [29] Dimorphic hemmorhoids(piles)
## [30] Heart attack
## [31] Varicose veins
## [32] Hypothyroidism
## [33] Hyperthyroidism
## [34] Hypoglycemia
## [35] Osteoarthristis
## [36] Arthritis
## [37] (vertigo) Paroymsal  Positional Vertigo
## [38] Acne
## [39] Urinary tract infection
```

```
## [40] Psoriasis
## [41] Impetigo
## 41 Levels: (vertigo) Paroymsal  Positional Vertigo Acne ... Varicose veins
```

Splitting the Data

We can now split the data into 70% training and 30% testing data. I used createDataPartition() function of library caret for random splitting resulting in balanced outcome classes.

```
## Training and Testing Dataset

Train_index <- createDataPartition(Disease$prognosis, p = .7, list = FALSE, times = 1)

train <- Disease[ Train_index,]
test <- Disease[-Train_index,]
```

E.Modelling and Results

After splitting the data into training and testing data, we will apply various machine learning algorithms to make various models, and compare them using various metrics.

Decision Tree

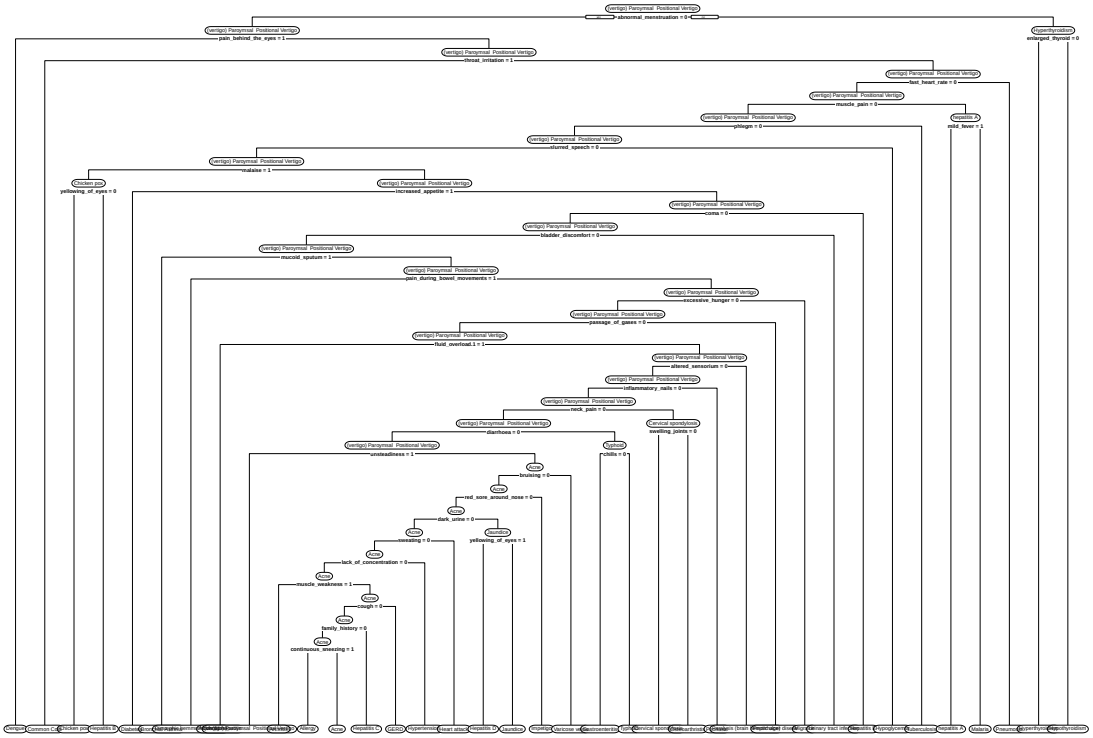
Decision Trees are a type of Supervised Machine Learning (that is you explain what the input is and what the corresponding output is in the training data) where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves. The leaves are the decisions or the final outcomes. And the decision nodes are where the data is split.

I trained the decision tree on the training dataset using all variables of symptoms as predictors and prognosis as response variable.We also did pruning of tree to make the better visualized tree.

```
Disease_tree <- rpart(prognosis ~ ., data = train, method = "class")



# Plot the tree
rpart.plot(Disease_tree, extra = 0, yesno = TRUE)
```

```
## Warning: All boxes will be white (the box.palette argument will be ignored) because
## the number of classes in the response 41 is greater than length(box.palette) 6.
## To silence this warning use box.palette=0 or trace=-1.
```

```
## Warning: labs do not fit even at cex 0.15, there may be some overplotting
```



```
# Make prediction on test data
pred1 <- predict(Disease_tree, newdata = test, type = "class")


CM_dt1 <- confusionMatrix(pred1, test$prognosis)
CM_dt1$overall
```

```
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper   AccuracyNull
##     0.85365854     0.85000000     0.83458595     0.87131132     0.02439024
## AccuracyPValue  McnemarPValue
##     0.00000000            NaN
```

```
# Extract the table from the confusion matrix
table <- as.table(CM_dt1$table)

## Sensitivity

# Calculate the sum of true positives
TP <- sum(diag(table))

# Calculate the sum of false negatives
FN <- sum(rowSums(table)) - TP

# Calculate Micro-averaged Recall
Micro_Recall <- TP / (TP + FN)

# Print Micro-averaged Recall
cat(paste("Sensitivity = ", Micro_Recall), "\n")
```

```
## Sensitivity =  0.853658536585366
```

```r
## Specificity

# Initialize sums
sum_TN <- 0
sum_FP <- 0

# Calculate the sum of true negatives and false positives for each class
for (i in 1:nrow(table)) {
  TN <- sum(table[-i, -i])
  FP <- sum(table[-i, i])
  sum_TN <- sum_TN + TN
  sum_FP <- sum_FP + FP
}

# Calculate Micro-averaged Specificity
Micro_Specificity <- sum_TN / (sum_TN + sum_FP)

cat(paste("Specificity = ", Micro_Specificity), "\n")
```

```
## Specificity =  0.996341463414634
```

Results

The decision tree model achieved the accuracy of 89.70%

Naive Bayes

```r
## Naive Bayes
library(e1071)

model <- naiveBayes(prognosis ~ ., data = train)
print("Prediction on Test data")
```

```
## [1] "Prediction on Test data"
```

```r
predictions <- predict(model, newdata = test)

CM_dt1 <- confusionMatrix(predictions, test$prognosis)
CM_dt1$overall
```

```
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper   AccuracyNull
##     1.00000000     1.00000000     0.99750388     1.00000000     0.02439024
## AccuracyPValue  McnemarPValue
##     0.00000000            NaN
```

```r
# Extract the table from the confusion matrix
table <- as.table(CM_dt1$table)

## Sensitivity
```

```r
# Calculate the sum of true positives
TP <- sum(diag(table))

# Calculate the sum of false negatives
FN <- sum(rowSums(table)) - TP

# Calculate Micro-averaged Recall
Micro_Recall <- TP / (TP + FN)

# Print Micro-averaged Recall
cat(paste("Sensitivity = ", Micro_Recall), "\n")
```

```
## Sensitivity =  1
```

```r
## Specificity

# Initialize sums
sum_TN <- 0
sum_FP <- 0

# Calculate the sum of true negatives and false positives for each class
for (i in 1:nrow(table)) {
  TN <- sum(table[-i, -i])
  FP <- sum(table[-i, i])
  sum_TN <- sum_TN + TN
  sum_FP <- sum_FP + FP
}

# Calculate Micro-averaged Specificity
Micro_Specificity <- sum_TN / (sum_TN + sum_FP)

cat(paste("Specificity = ", Micro_Specificity), "\n")
```

```
## Specificity =  1
```

Results

The Naive Bayes model has achieved an accuracy of 1 which is far better than that of decision tree model.

Support Vector Machine

```r
### Support Vector Machine

model_svm <- svm(prognosis ~ ., data = train)

predictions_svm <- predict(model_svm, newdata = test)

CM_dt1 <- confusionMatrix(predictions_svm, test$prognosis)
CM_dt1$overall
```

```
##       Accuracy          Kappa  AccuracyLower  AccuracyUpper   AccuracyNull
##     1.00000000     1.00000000     0.99750388     1.00000000     0.02439024
## AccuracyPValue  McnemarPValue
##     0.00000000            NaN
```

```r
# Extract the table from the confusion matrix
table <- as.table(CM_dt1$table)

## Sensitivity

# Calculate the sum of true positives
TP <- sum(diag(table))

# Calculate the sum of false negatives
FN <- sum(rowSums(table)) - TP

# Calculate Micro-averaged Recall
Micro_Recall <- TP / (TP + FN)

# Print Micro-averaged Recall
cat(paste("Sensitivity = ", Micro_Recall), "\n")
```

```
## Sensitivity =  1
```

```r
## Specificity

# Initialize sums
sum_TN <- 0
sum_FP <- 0

# Calculate the sum of true negatives and false positives for each class
for (i in 1:nrow(table)) {
  TN <- sum(table[-i, -i])
  FP <- sum(table[-i, i])
  sum_TN <- sum_TN + TN
  sum_FP <- sum_FP + FP
}

# Calculate Micro-averaged Specificity
Micro_Specificity <- sum_TN / (sum_TN + sum_FP)

cat(paste("Specificity = ", Micro_Specificity), "\n")
```

```
## Specificity =  1
```

Results of Support Vector Machine

The SVM model also achieved accuracy of 1 which is same as Naive Bayes

F. Summary

After, comparing the models, it is shown that SVM and Naive Bayes are better performing than Decision tree in terms of accuracy, specificity and sensitivity.

However, in real world scenario, it is not practically possible to a achieve an accuracy of 1. So it is derived that the models are overfitting due to unavailabilty of accurate data.