Diagnosis of Mesothelioma's Disease Based on Classification

Sahba Akhavan[1], Shaojun Zhang[1], Delaram Ghoreishi[2]

[1]Department of Statistics, [2]Department of Physics, University of Florida

April 14, 2016

Malignant mesothelioma's (MM)

`Mesothelioma` is a type of cancer that develops from the thin layer of tissue covering many of the internal organs.
Lining of the lungs and chest wall is the most common area affected. It could also affect sac of the heart or lining of the abdomen.

Malignant mesothelioma's (MM)

`Mesothelioma` is a type of cancer that develops from the thin layer of tissue covering many of the internal organs.
Lining of the lungs and chest wall is the most common area affected. It could also affect sac of the heart or lining of the abdomen.
Symptoms include:

- Shotness of breath due to fluid around the lung (pleural effusion).
- Chest wall pain.
- cough and ...

Symptoms appear slowly which makes the diagnosis hard, and they don't immidiately alert the doctor to diagnose MM.
Main cause of Mesothelioma is prolonged inhalation of `asbestos` fibres.
Asbestos is a naturally occuring silicate mineral.

Diagnosis of MM based on Classification

Data set is from patient's hospital reports from Dicle University, Faculty of Medicine's, Turkey. It consists of symptoms of 324 MM patients.
Each variable has 34 features.
Our task is to diagnose MM using classification techniques.
We excluded the diagnosis.methods from our data set because of its complete resemblance to class.of.diagnosis.

Diagnosis of MM based on Classification

Data set is from patient's hospital reports from Dicle University, Faculty of Medicine's, Turkey. It consists of symptoms of 324 MM patients.
Each variable has 34 features.
Our task is to diagnose MM using classification techniques.
We excluded the diagnosis.methods from our data set because of its complete resemblance to class.of.diagnosis.

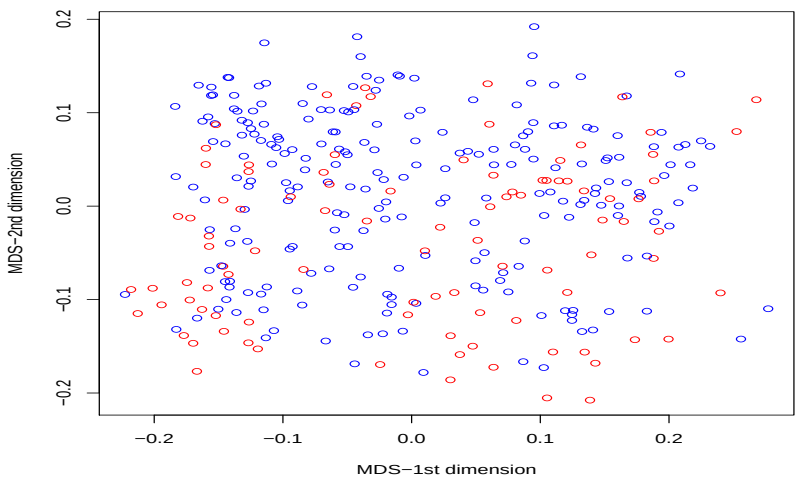

Figure 1 : MDS plot of the data set after removing diagnosis.methods.

Results

Among the methods used RF gave the best result.

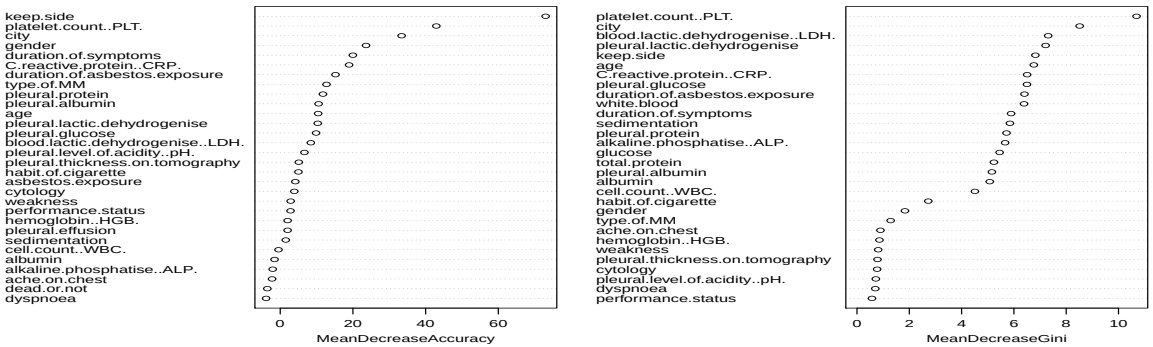

Figure 2 : varivable importance plot for RF

Method	**Error Rate**
LDA	0.25
Tree	0.28
KNN	0.33
SVM	0.30
RF	0.18

References

O. Er, A. C. Tanrikulu, A. Abakay, and F. Temurtas.

An approach based on probabilistic neural network for diagnosis of Mesothelioma's disease.

Comput. Electr. Eng., 38(1):75–81, jan 2012.

George Michailidis.

Applied multivariate analysis.

J. Econom., 3(3):320, 1975.

Questions

Thank You.