**Predicting Scaled Sound pressure Using Bayesian Network**

**Abasiekeme Attang**

In this case study I am using Bayesian Networks to predict Scaled Sound Pressure.

**ABSTRACT**

A Bayesian network (from now on BN) of a set of variables X = {X1, . . . , Xn} represents a joint probability distribution over those variables. It consists of a network structure that encodes assertions of conditional independence in the distribution and a set of conditional probability distributions corresponding to that structure. It is graphically represented by directed acyclic graphs, whose nodes denotes the random variables, which may be observable quantities, latent variables, unknown parameters or hypotheses. Edges represent conditional dependencies, so that nodes which are disconnected represent variables which are conditionally independent of each other.

Bayesian networks (BNs) are a type of graphical model that encode the conditional probability between different learning variables in a directed acyclic graph. There are benefits to using BNs compared to other unsupervised machine learning techniques. A few of these benefits are:

1. It is easy to exploit expert knowledge in BN models.
2. BN models have been found to be very robust in the sense of i) noisy data, ii) missing data and iii) sparse data.
3. Unlike many machine learning models (including Artificial Neural Network), which usually appear as a “black box,” all the parameters in BNs have an understandable semantic interpretation.

Keywords

* Bayesian networks; Scaled Sound Pressure.

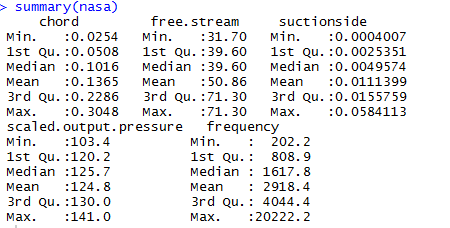
**DATASET**

The NASA data set comprises different size NACA 0012 airfoils at various wind tunnel speeds and angles of attack. The span of the airfoil and the observer position were the same in all of the experiments. It uses the following attributes

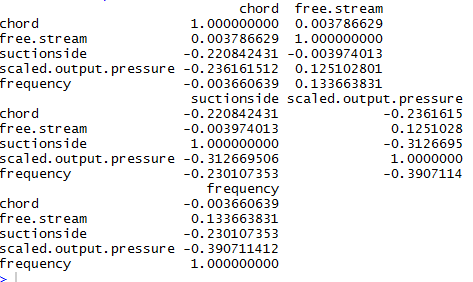
* Frequency, in Hertz’s.
* Angle of attack, in degrees.
* Chord length, in meters.
* Free-stream velocity, in meters per second.
* Suction side displacement thickness, in meters.
* Scaled sound pressure level, in decibels.

**Exploratory Data Analysis**

The data collected from UCI Irvine machine learning dataset. The NASA datasets. The diagram below shows the summary of the data which shows the minimum, median,mean,maximum,median and quartiles of the all the variables involved.



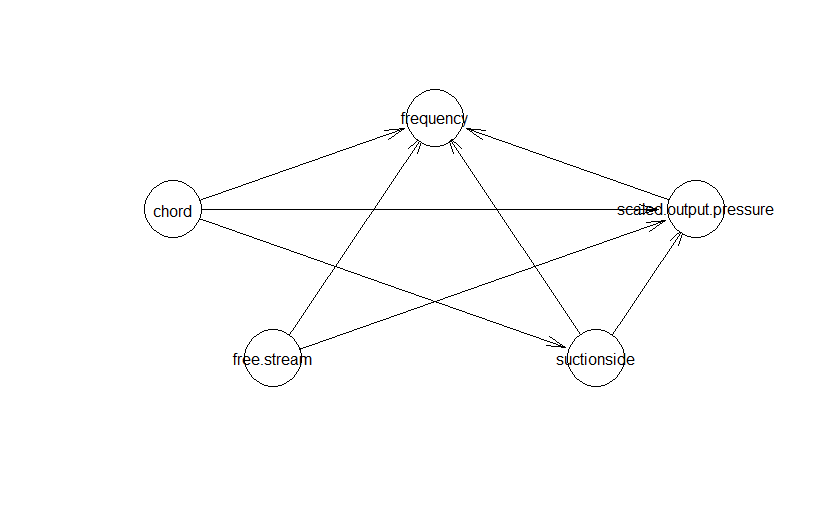
From the plot we see they are no strong correlations between the variables.



**Experimental Results**

The analysis performed on the data I had to use to real numbers in order to perform Bayesian networks. I used R to perform this and used the library “bnlearn” to implement it, I used the bnlearn package in R to learn the structure of my Bayesian network and its parameter .The plan is to predict the value of a node (Scaled sound pressure) given the value of other nodes as evidence.

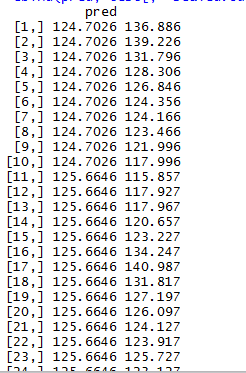
The plot below shows Bayesian Network structure on training set data using hill climbing module of bnlearn.



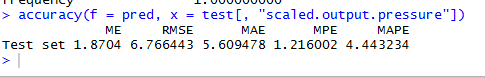
**Experimental Analysis**

Based on the Bayesian Network performed on the data, I determined the predicted scaled output pressure and compared it with the real scaled output pressure and checked the accuracy of the implementation of Bayesian Network.

The diagram below shows the part of the predicted values and the actual value.



The diagram below shows the accuracy of the Bayesian Network.It shows the Mean Error(ME) ,Root Mean Square Error(RMSE), Mean Absolute Percentage Error(MAPE) ,Mean Percentage Error(MPE) values which shows the goodness of fit of this specific Bayesian Network model.



**Conclusion**

The case study presented here have demonstrated some of the potential uses for Bayesian approaches in predicting scaled sound pressure and I observed that the accuracy was relatively high and predictions were close to the range. I observed from this case study when used in conjunction with statistical techniques, the graphical model has several advantages for data modeling. Bayesian network is one of efficient tools in the data-mining, dealing with incomplete data and describing the relativities of data in quantity and the correlation in graphics methods, applied to data mining can be reflected, the potential dependent relationship between variables.