Study of Machine Learning Algorithms

Abstract

A study of machine learning algorithms and their implementation on the cloud. Machine learning algorithms surveyed include supervised learning techniques like linear regression models, neural networks, fuzzy neural networks, decision trees , support vector machines, deep neural networks.. Fuzzy rule based inference systems, and algorithms that extract fuzzy rules from data are also surveyed here. The second part of the work involves understanding how basic machine learning algorithms can be implemented in the cloud environment on scalable data. Finally, these algorithms will be tested against a residential energy consumption data-set and bench marked for capability of each of the models to predict energy consumption based a set of attributes.

Introduction

Supervised learning is a branch of machine learning where a statistical model is generated based on examples. Here attributes of data which has to be predicted is used to train a model by giving examples of input and observed output. Once the model *learns* how to generate the output under an acceptable error rate, this model is tested using another set of data where only input is given and the predicted output is tested against observed output and the error is calculated. A model is able to predict the output by different methods which depends on the technique used to created the model. For example in a neural network the weights of the network is altered based on techniques like back propagation using gradient descent to reduce an error function, thus generating a model that gives a *predicted* output from a given set of inputs.

There is another set of statistical techniques and other techniques known as unsupervised learning. Where a set of data is analyzed by algorithms automatically to understand the structure and underling patterns of the data . Here the optimization or the number of iterations is not based on minimizing an error function but based in creating sets of similar data.

Fuzzy Logic

Fuzzy logic is a form of logic which rather than based on typical two-valued logic tries to model different phenomenon using a non-crisp logic which can have any value between 0 and 1, this is called the truth value. Lofti Zadeh proposed this new form of logic in his 1965 paper on *fuzzy set theory.*

A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership (characteristic) function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets, and various properties of these notions in the context of fuzzy sets are established. In particular, a separation theorem for convex fuzzy sets is proved without requiring that the fuzzy sets be disjoint.[ZAD65]. In the study I did I used a fuzzy rule based inference system with Wang and Mendel model[WM92] to solve a regression task. The implication function used

is Zadeh implication function with 7 linguistic variables and a Gaussian membership function.

A number of automatic techniques has been proposed to generate these fuzzy rules from numerical data. One of the most interesting family of techniques, due to its simplicity and quickness, is the ad-hoc data-driven methods.[DECSAI] .

Wang and Mendel Model

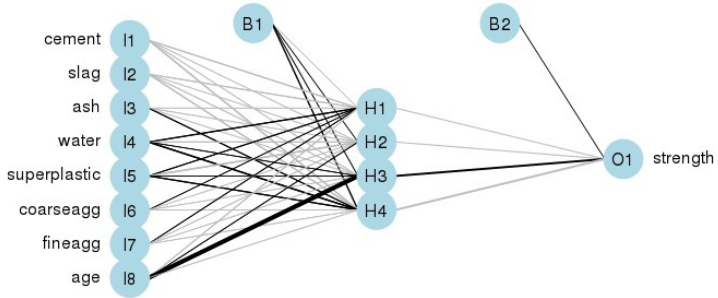
In 1992, Wang and Mendel (Wang and Mendel, 1992) proposed a new approach to generate fuzzy IF-THEN rules from numerical data pairs, now known as the Wang-Mendel (WM) method. The generated fuzzy rule base can be combined with experts' knowledge to construct a combined system of rules. They proved that the resulting fuzzy system is capable of approximating any nonlinear continuous function on a compact set to arbitrary accuracy.[SPE13]

In Wang and Mendel method is one of the basic fuzzy rule learning method for function approximation. Here based on predefined grid of X and Y, it determines best possible output fuzzy sets for each rule described using the fuzzy grid. Running this algorithm through a dataset, it gets clustered and a fuzzy value is assigned to corresponding fuzzy set.

Neural Networks

A neural network is a "...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs. [RHN89]

The network is composed of a large number of highly interconnected processing elements (neurons) working in parallel to solve a specific problem. Neural networks learn by example.[NN]. A special type of neural network is the Multi-layer perceptron which is one of the most popular neural network architectures that fit a wide range of applications such as forecasting, process modeling and pattern classification. MLP are feed-forward neural networks trained with a technique called back propagation. It has 3 layers, one input layer, one hidden and an output layer. These layers are connected using vectors of connection weights. An activation function is applied on the summed up weights on input values. The weights are re-calculated using the back propagation adjusting them as needed so that relationship between input and output is established . Once this is established this network can be used for predicting outputs from input values.

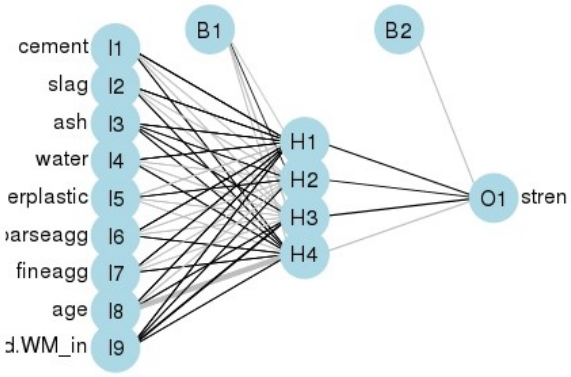


A simple neural network

This neural network was created using a dataset concrete which has attributes like strength, cement, slag, ash, water, superplastic, age etc and tries to predict the strength of concrete formed using these attributes. Here there are three layers, the input layer, a hidden layer and the output layer. Here the input layer has 9 nodes that corresponds to the number of input attributes, 4 hidden nodes and one output node. There are two biases which are used to bias the network so that it does not saturate.

I designed this neural network in R, and the same dataset was used to created a FRBS, or a fuzzy rule based inference system, another network with one more input that these attributes, which is the output of the FRBS system. Then I calculated the correlation value based on predicted and observed output, the strength of concrete. The results show that the system that combines, the fuzzy inference system with a neural network gives the best correlation among the models.

The final network from became



The calculated Correlation table is:

corr.neuro	corr.fuzzy	**corr.neuro.fuzzy**
[1,] 0.7202496	0.5480245	**0.9314786**

The calculated MSE table is:

MSE.nnet	MSE.fuzzy	**MSE.neuro_fuzzy**

0.012659327 0.034375753 **0.006202421**

As we can see from the tables above the correlation and mean squared value for the model that uses both fuzzy and neural network is the best among the three models.

Further on:

Tree Models

SVM

Deep Neural Networks

References

ZAD65: Lotfi A. Zadeh, Information and Control , 1965
WM92: Wang, L.-X. ; Dept. of Electr. Eng. & Comput. Sci., California Univ., Berkeley, CA, USA ; Mendel, J.M., Neural Networks, IEEE Transactions on , 1992
DECSAI: J Casillas, O. Cord´on, F. Herrera, Improving the Wang and Mendel'sFuzzy Rule Learning Method byInducing Cooperation Among Rules, 2011
SPE13: Minshen Hao, Jerry M. Mendel, Extracting IF-THEN Rules from Numerical Data using Wang-Mendel Methods , 2011
RHN89: Maureen Caudill, Dr. Robert Hecht-Nielsen, Neural Network Primer: Part I" by Maureen Caudill, 1989
NN: Ms. Sonali. B. Maind, Research Paper on Basic of Artificial Neural Network,