



**HAL**  
open science

# Mathematical models of classification algorithm of Machine learning

Dutta Nabanita, Subramaniam Umashankar, Padmanaban Sanjeevikumar

## ► To cite this version:

Dutta Nabanita, Subramaniam Umashankar, Padmanaban Sanjeevikumar. Mathematical models of classification algorithm of Machine learning. International Meeting on Advanced Technologies in Energy and Electrical Engineering (IMAT3E'18), Nov 2018, Fez, Morocco. <hal-01941420v2>

**HAL Id: hal-01941420**

**<https://hal.science/hal-01941420v2>**

Submitted on 14 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**Mathematical models of classification algorithm of Machine learning**

<sup>1</sup>Nabanita Dutta, nabajhulikbarbi@gmail.com <sup>2\*</sup> Umashankar.S, shankarums@gmail.com  
<sup>2</sup> Sanjeevikumar P, san@et.aau.dk

<sup>1</sup>lead presenter, Nabanita Dutta

<sup>1</sup>, nabajhulikbarbi@gmail.com, VIT Vellore, India

<sup>2\*</sup> Prince Sultan university, Riyadh, Saudi Arabia

<sup>3</sup> Aalborg University, Denmark

**Abstract**

Machine learning algorithm has brought the augmenting change in the field of artificial intelligence which espoused human discerning power in a splendid manner. The algorithm has various categories among which classification is the most popular part. Support vector machine algorithm, logistic regression, naïve bays algorithm, decision tree, boosted tree, random forest and k nearest neighbour algorithm all are under classification algorithms. Classification process needs some pre-defined method which leads the method for choosing the train data from the sample data given by the user. Decision making is the heart of any classification algorithm as supervised learning stands out on the decision of users. So the strong mathematical model based on conditional probability lies behind each algorithm. This paper is the study of those mathematical models and logic behind various classification algorithms which help to create a strong decision for users to make the training dataset based on which machine can predict the proper output.

**Markov chain model**

Markov chain model is a statistical and mathematical setup which has some hidden layers and can be represented as the simple Bayesian network which is directly visible to the observer. This model has a remarkable contribution in the field of supervised and reinforcement learning and for pattern recognition. If any example is considered for two classes i.e. A and B and it has 4 transitions when the system is in A then it can be considered as a transition of B similarly when a system is in B it can be considered as a transition of A (Fig1). In that way, a transition matrix will be formed which will define the probability of transition of the state. In that way not only in two classes but also with n no of classes or states the model can be built up.

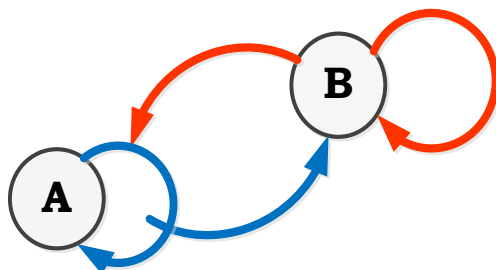


Fig 1: Two state Markov chain model

So the Markov chain model is based on conditional probability. The main target of this structure is to concentrate on the completion of a task is the best-suited way which requires the probability of the summation of all sequences in its best-suited way. The probability of the observant in a particular sequence has been described in the equations (Eqn 1 to 4).



International Meeting on Advanced Technologies in Energy and Electrical Engineering

$$Y = y(0), y(1), \dots, y(L-1) \tag{1}$$

The length of the L is defined as

$$P(Y) = \sum_x P(Y | X)P(X) \tag{2}$$

The sum of the overall hidden node sequences

$$X = x(0), x(1), \dots, x(L-1) \tag{3}$$

Using the chain rule it is defined that

$$P(e_i)P(e_2 | e_1)P(e_3 | e_1, e_2)P(e_4 | e_1, e_2, e_3)KP(e_i | e_1, e_2, Ke_{i-1}) \tag{4}$$

The Markov chain assumption depends on  $e_i$ .

**Example of Markov chain model**

In the heavy industry like oil and gas industry, petrochemical industry, in power plant moreover in household purposes and in agriculture, the pumping system is a very instigate part so the health of the pumping system is needed to be under proper vigilance. Machine learning algorithm has conveyed the bewildering elucidation where both incessant monitoring and ground level fault detection are conceivable. Various faults occur in the pumping system where cavitation is one of the faults among these [1]. Cavitation occurs when the bubbles are created in the area of the impeller and sudden pressure drops. So detection of cavitation in the pumping system is a challenging task. If it is considered that total points collectively with faulty and no faulty condition is 300 and only cavitation fault points are 240, then the implementation of Markov model based on conditional probability is described in the equations ( Eqn 5-7).

$$P(\text{Trained points} | \text{Faulty points (cavitation)}) = \frac{P(\text{Faultypoint s(cavitation)})}{P(\text{totaltrainedpoint s})} \tag{5}$$

$$= \frac{C(\text{Faultypoint s(cavitation)})}{C(\text{totaltrainedpo int s})} \tag{6}$$

$$= \frac{240}{300} \tag{7}$$

$$= 0.8$$

So the choosing of faulty points is more preferable when the user can define the faulty points properly.

**The mathematics behind the algorithms**

Each and every classification algorithm is built up with strong mathematical models and logic. Though the decision tree is one of the classification algorithms every algorithm is bound to make one decision before starting of making the trained model. The logic of the SVM is for finding out the optimal separating hyperplane that maximizes the margin of training data [2]. If two classes A and B are there, for separating A and B a hyperplane is required which is called separating hyperplane, so choosing of the hyperplane is very vital. If hyperplane is wrongly chosen then it is very difficult for the classification of two classes [3] (Fig 2). The hyperplane is constrained by the equation 8.

$$\frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i + C \sum_{i=1}^N \xi_i^* \tag{8}$$

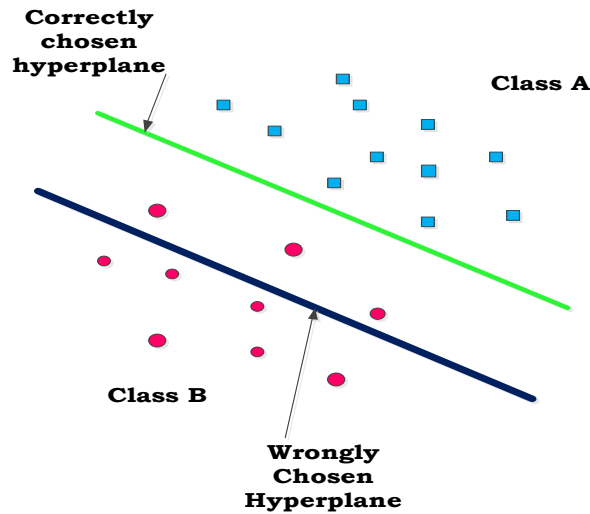


Fig 2: Choosing of hyperplane for classification in SVM

Similarly, k nearest neighbour algorithm needs the proper distance from the testing point to the classification point. The distance  $d$  has a relationship between test point  $y$  with classification point  $x$  [4] (eqn 9).

$$d_{st}^2 = (x_s - y_t)^{V-1} (x_s - y_t)' \tag{9}$$

Similarly decision tree algorithm needs proper decision-making tree which will take the decision for predicting the best-suited output, naive bays algorithm stands on the theory of conditional probability (eqn 4), logistic regression needs continuous data for analysis and prediction of output is based on sigmoid function (eqn 10), random forest stands on the theory of Chebychev inequality combination of mean and standard deviation (eqn 11).

$$h(x) = \frac{1}{1 + e^{-(\theta_0 + \theta_1 * X)}} \tag{10}$$

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \tag{11}$$

**Conclusion:**

The article presents the Markov chain model based on conditional probability which describes the classification algorithms for prediction of faulty points in the system accurately.

**References:**

[1].Nabanita Dutta, Umashankar.S, V.K.Arun, P. Sanjeevikumar, Zbigniew Leonowicz, P.Wheeler, "Centrifugal Pump Cavitation Detection Using Machine Learning Algorithm Technique" 18th International Conf. on Environmental and Electrical Engg., IEEE-I&CPS/IEEE-EEEIC'18, Paleramo, Italy.

[2].A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, 2007.

[3].L. Alfayez, D. Mba, and G. Dyson, "The application of acoustic emission for detecting incipient cavitation and the best efficiency point of a 60 kW centrifugal pump: the Case study," *NDT E Int.*, vol. 38, no. 5, pp. 354–358, 2005

[4].G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN Model-Based Approach in Classification."