Week 4 Report: Backpropagation review

Lam Dang

2017-04-27

1 Gradient Descent

Gradient Descent is a popular method in Data Mining to obtain the most suitable model. Gradient Descent usually used in tandem with other algorithm in order to create a model for prediction or classification, such as Linear Regression, Neural Network etc.

The concept of Gradient Descent is to "go down" the gradient "well" at a predefined rate. After each "step" the algorithm re-evaluate for new weight vector. These steps are repeated until a pre-defined number of iterations or until the global minima have been reach.

Another view on the algorithm is to consider it as an optimization process [1]. In this view, Gradient Descent is considered as a constrained minimization problem.

1.1 Mathematical Model

Bishop [2, p240] provide a simple formular to update weights vector base on Gradient information

$$w^{(\tau+1)} = w^{(\tau)} - \eta \nabla E(w^{(\tau)}) \tag{1}$$

whereas η is the pre defined learning rate and $E(w^{(\tau)})$ is the loss function defined in the following section.

1.2 Loss Function

Bishop [2, p242] also describe an evaluation of weight vector base on derivative of error function.

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} ||y(x_n, w) - t_n||^2 \tag{2}$$

2 Backpropagation

Backpropagation is one of the the dominating method for learning with Artificial Neural Network. The concept of Backpropagation is relatively simple however they can provide robust learning scheme and models. [1]

2.1 Algorithm

Backpropagation algorithm use indefinitely differentiable activation function. We denote these function as F and their derivative as F'. The algorithm went through 2 steps

2.1.1 Forward Propagation

The input X_i is feed into the network. The functions $F(X_i)$ are calculated and propagate forward into next layers. The derivative functions $F'(X_i)$ are stored

2.1.2 Back Propagation

The

2.2 Weight update

2.3 Learning Rate

3 Current Result

3.1 Linear Regression

3.2 Extreme Learning Machine

3.3 Backpropagation

References

[1] Yann Lecun. A theoretical framework for back-propagation. In *Artificial neural networks*. IEEE Computer Society Press, 1992.

[2] Christopher M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, 2006.

[3] Ral Rojas. *Neural Networks*. Springer Berlin Heidelberg, 1996. DOI: 10.1007/978-3-642-61068-4.