

Development of An Optimal Neural Network For Avalanche Forecast In Himalayan Region

V.Rihani

Head, Dept. of E&E, PEC
Sector-12, Chandigarh
vrihani@yahoo.com

Mahesh Bansal

Student of ME, PEC
Sector-12, Chandigarh.
mahesh_bansal@sify.com

Atul Parti

Student of ME (CSE & IT), PEC
Sector-12, Chandigarh.
atulparti@yahoo.com

Rashpal Kaur

Scientist-“C”
Snow & Avalanche Study Establishment
rashpalkaur@yahoo.com

Abstract - *This paper deals with the application of a well-known data mining technique, multi-layer back-propagation neural network, for forecasting of an avalanche in Himalayan region. Meteorological and snow data for Himalayan region has been used for training the neural network. EasyNN-plus 6.0g, neural network software for Microsoft windows, is used for the development of an optimal neural network-PUSHPDEV. The system tries to model the decision making process of a pragmatic expert. PUSHPDEV can forecast whether an avalanche will trigger on a particular day from November to April. The network accepts eighteen inputs and produces an output whose value is zero or one, zero for no avalanche and one for avalanche on that day.*

Keywords: Avalanche, data mining, data warehouse

1 Introduction

An avalanche is a mass of snow or ice that slides down the side of a mountain rapidly. Avalanches [1] are generated by structural weakness in the snowcover. Avalanches occur due to the failure of the snowpack to withstand the stresses placed upon it by its self weight and other imposed forces. The forecasting factors of avalanche are divided into three classes [2-3]. The higher the class number less relevance is the data for avalanche formation. These classes are:

Class I: Stability factors, including current avalanche activity, stability tests (e.g. Rutschblock, Shovel test) as well as other signs of instability, such as cracking of the snowcover. This is most relevant data for avalanche formation.

Class II: Snowpack factors, such as the structure of the snowpack, including the existence of weak layers, crystal forms and sizes, densities and snow

temperatures. This data has secondary relevance in avalanche formation.

Class III: Meteorological factors, such as the temperature, winds, precipitation and humidity which bears indirect relevance to avalanche formation.

Primarily avalanche forecasting [4] is done by an expert, acting on their knowledge and experience with snow stability situations. The physical [5] principles underlying avalanche formation remain an area of active research as they are not yet fully understood. A forecasting tool can therefore be used as an aid to expert in decision-making for prediction of an avalanche. Different types of techniques are used in avalanche forecast. They can be categorized as non-computer assisted [6] and computer assisted techniques [4-5][7-17].

Avalanches in Himalayan are a serious threat to both human as well as to property. Avalanches results in closure of roads which affects the necessary supplies and services to both army and humans. Fuzzy rule bases system [14] and nearest neighbour method [17] are two computer assisted tools, used for avalanche forecasting in Himalayan region. The need for development of another tool is to use it along with already existing tools so as to have a better avalanche warning system for Himalayan region.

Data mining [18] is known as “knowledge discovery in databases (KDD)”. It is the process of extraction of patterns representing knowledge stored in either data warehouse or databases. Data mining is a multidisciplinary field, drawing work from area including database technology, artificial intelligence, neural network, machine learning, statistics, knowledge acquisition, high performance computing, information retrieval, and data visualization. C4.5, C5, CART and feedforward backpropagation are different types of data mining algorithm used. Artificial neural

network (feedforward backpropagation), a form of data mining, has the advantage that it can handle inconsistent data. Artificial neural network can therefore is a good choice for forecasting problems like avalanche where data collected over years is inconsistent. Artificial neural network is an attempt to mimic the biological decision making process of the human. Artificial neural network [18] development is a two-step process. In the first step, supervised type of learning is used to build a model. In step two, the model is used for the classification and the accuracy is estimated. If estimated accuracy of the model is acceptable then the model is used for prediction.

Artificial neural networks have been developed in south central Colorado of United States [4]. They uses fifteen metrological and snow parameters, and output is an index whose value ranged from zero to one, with higher values corresponding to increased likelihood's of avalanche activity. These models have the accuracy of predicting avalanche activity up to 78 to 91 percent at different learning rate.

EasyNN-plus 6.0g, is neural network software for Microsoft windows, is used for the development of an optimal neural network. All the component parts of EasyNN-plus are implemented as C++ reusable classes. The inspiration to use EasyNN-plus was based on the fact that it has a graphical user interface and backpropagation algorithm. EasyNN-plus uses double linked lists for implementation to store the examples, the nodes and the connections. The lists can be processed quickly in both directions simultaneously, and it can also be extended and contracted dynamically.

Table 1 Data used in PUSHPDEV

	1999	2001	2002	2003	2004	2005
Jan	√	√	√	√	√	*
Feb	√	√	√	√	√	*
Mar	√	√	√	√	√	*
April	√	√	√	√	√	*
Nov	√	√	√	√	*	
Dec	√	√	√	√	*	
√ : Data used for training * : Data used for testing						

2 Data

The present study lies in Lower Himalayan Zone and data for a particular site, is provided by SASE (Snow & Avalanche Study Establishment). It includes metrological and snow data, which falls in Class III, least relevant data for avalanche forecast. Data available for different years is shown in Table 1. Following is the description of data.

1) Daily weather data, which records meteorological and snow parameter like minimum temperature, maximum temperature, dry temperature, average wind

speed, fresh snow amount, fresh duration and standing snow etc. These parameters are taken twice in a day, at 8:30 and 17:30 hours.

2) Avalanche occurrence data, which specify when an avalanche occurred.

3 Method

The development of an optimal neural network has three main processes: objective determination, data preparation and neural network modeling. Detail descriptions of these processes are

3.1 Objective Determination

The basic aim of this process is to obtain clear understanding of the model to be developed. Objective determination leads to the actual specification to be used for the development of the model. The output of this phase is the problem definition and expected result. Problem definition for our work is *"To develop an optimal neural network model to predict whether an avalanche will trigger on a particular day or not in Himalayan region."* The expected output of the model is a value, which is zero or one, zero for no avalanche and one for avalanche on that day.

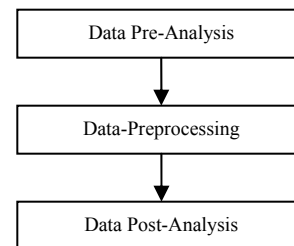


Figure 1 Data preparation scheme

3.2 Data Preparation

Data preparation is an important and critical step in neural network modeling. Data preparation includes all operation that is carried out on data for development of the model. Data preparation [19] includes data pre-analysis, data preprocessing and data post-analysis. Data preparation scheme is shown in Figure 1.

3.2.1 Data Pre-Analysis

Data pre-analysis phase includes collection of metrological and snow data for last few years. It also identifies main data variables from the collected data, which was used later for the development of the model. The selections of variables were based on their importance to problem and quantity of data available. The parameters identified and used as input to PUSHPDEV are shown in Table 2.

Table 2 Eighteen input to the PUSHPDEV

Sr. No.	Input
1	Max temperature
2	Min temperature
3	Dry temperature
4	Dry temperature difference
5	Average maximum(Average of last three reading)
6	Average minimum(Average of last three reading)
7	Difference of average maximum(Difference of current and last average maximum temperature)
8	Difference of average minimum(Difference of current and last average minimum temperature)
9	Pressure difference
10	Average speed
11	Difference of average speed
12	Fresh snow continuous (sum of current and last fresh snow amount)
13	Fresh snow duration
14	Fresh snow water equivalent ($\log(1+x)/\log 181$, where x is the value of Fresh snow water equivalent).
15	Standing Snow
16	Difference in snow temperature
17	Snow Penetration below Surface(Is the depth of snowpack up to which the ram penetrometer penetrates with its own weight)
18	Characteristic of weather during proceeding one hour (new code developed by Code conversion method describes later.)

3.2.2 Data Preprocessing

Data pre-analysis step was followed by data preprocessing. Data preprocessing is used to extract key features from data and transforming data in the form that neural network can accept. Data preprocessing improves the quality of data, which influence the neural network models to be built. High quality data improves performance and learning time for the network. Different types of operations were performed on data to yield good results. These operations include substituting data for missing values, taking average, difference of values/average and log of values for different parameters. A new method called “*Code-Conversion*” is proposed, for the use of non-numeric data in artificial neural network. Non-numeric data used in our case is “*characteristic of weather during proceeding one hour*”, is the weather code assigned for different weather condition. Their impact on avalanche is not known, but they play an important role in the avalanche to trigger. The proposed method calculates the number of times each code exists and the number of times avalanche has triggered for that code, and then their ratio is calculated. This ratio is called “*Conversion Ratio*”. *New codes are assigned corresponding to the conversion ratio that is higher conversion ratio has higher new code.* The idea of introducing new code is to eliminate longer learning time and improve performance of trained network.

3.2.3 Post Analysis

Last step in data preparation is post analysis, which include dividing data into two parts, one for training and other for testing to verify the network performance. Data divided for training (including 99 records randomly selected for validation) and testing (for an entire season) is shown in Table 1.

3.3 Neural Network Modeling

Neural network modeling is an iterative process. It includes selecting an initial configuration. Then a number of experiments with each configuration are performed and the best network is retained. On each experiment, if under-learning occurs, then new network is tried by adding more neurons to the hidden layer/layers or by adding an extra hidden layer. If over-learning occurs, then new network is tried by removing hidden units. After training the network under predefined average error, network performance is verified against data kept for testing. If network performance against test data is satisfied, then the network is selected.

Table 3 Specification of PUSHPDEV

Number of input neurons	18
Number of output neuron	1
Number of neuron in first hidden layer	11
Number of neuron in second hidden layer	8
Learning rate	0.35
Momentum	0.35
Average training error	2%
Average validation error recorded	5.20%
Validating (99 random example)	96.00%
Number of cycles	3121
Accuracy of model for a season	87.38%

4 Main Result

Models were developed by using multi-layer (feedforward) backpropagation algorithm, with neurons on each layer fully connected to the neurons on the following layers. Various architectures were considered and selected model's specifications are shown in Table 3. Accuracy of *PUSHPDEV* is defined as the ratio of correct predictions against total number of predictions made. It shows 87.38% overall accuracy for an entire season i.e. 1st Nov, 2004 to 19th April, 2005.

5 Conclusions

The present study shows encouraging results for using artificial neural network to assist an expert in

prediction of an avalanche. The performance of neural network depends upon data used, which in our case falls in Class-III, is the least relevant data for avalanche forecast. Further artificial neural network can be developed by using stratigraphy data of the snowcover, which falls in Class II. However, it is important to realize that one particular model is not suited to all areas on the globe. This is especially true when attempting to model avalanche forecasts in the Himalayan region. The model developed is for a particular site in Himalayan region and this model may not be a good choice for forecasting of an avalanche in all parts of Himalayan region. For different location/site new model should be developed based on their historical data. The performance of these models will entirely depends upon the quality and quantity of data collected over the period of time. As each model has its own advantages, so hybrid models can be developed to incorporate advantages of more than one model.

6 Acknowledgement

We thank Dr. R.N.Sarwade, Director, Scientist-'G', Snow & Avalanche Study Establishment, Him Parisar, Plot No. 1, Sec-37-A, Chandigarh. We also like to thanks A.Ganju, joint Director, Scientist-'F', Snow & Avalanche Study Establishment. Last but not the least all the technical staff of Snow & Avalanche Study Establishment, whose help and support at all time resulted in the development of the project.

7 References

- [1] LaChapelle, E.R., "Avalanche forecasting- A modern synthesis". Int Association Science Hydrology, Pub. 69, pp. 75-84, 1965.
- [2] McClung, D.M, Schaerer, P, 1993: The Avalanche Handbook: The Mountaineers, Seattle, USA.
- [3] Zeidler, A. and Jamieson, B., "Computer Assisted Avalanche Forecasting: Skier-Triggered Avalanche", Paper presented Western Snow Conference, Vancouver, 2004.
- [4] Stephens, J., Adams, E., Huo, X., Hicks, J., and McCarty, D., "Use of neural networks in avalanche hazard forecasting", Proceedings - International Snow Science Workshop, 30 October-3 November, Snowbird, Utah, pp. 327-340, 1994.
- [5] Joseph, S.W., Lakeman, G.P., "An expert system to support snow avalanche forecasting", IEEE International Conference on System, Man and Cybernetics (3), pp. 22-25, 1995.
- [6] Fohn, P.M.B., "The 'Rutschblock' as a practical tool for slope stability evaluation". The Avalanche Review, Vol 7, No. 6, 1989.
- [7] Barbolini, M. and Keylock, C.J., "A new method for avalanche hazard mapping using a combination of statistical and deterministic models", Natural Hazards and Earth System Sciences (2002) 2: 239-245, 2002.
- [8] Lehning, M., Bartelt, P., Brown B., Fierz, C., and Satyawali, P., "A physical SNOWPACK model for the Swiss avalanche warning Part II. Snow microstructure", Cold Regions Science and Technology 35, 147- 167, 2002.
- [9] Leprettre, B., Navarre, J.P., Taillefer, A., Danielou, Y., Panel, J.M and Touvier, F., "Reliable Estimation of Avalanche Activity Using Seismic Methods", ISSW Proceeding, 1996.
- [10] Zeidler, A. and Jamieson, B. "Estimating the strength of faceted snow layers for an avalanche forecasting model", Proceeding of the International Snow workshop in Penticton, British Columbia, 29-sept to 4 oct, 2002.
- [11] Davis, R. E., K. Elder, D. Howlett and E. Bouzaglou, "Storm and weather factors related to dry slab avalanche activity", Proceedings - International Snow Science Workshop 1998, pp.25-34, 1998.
- [12] Davis, R. E., and Elder, K. "Application of Classification and Regression Trees: Selection of Avalanche Activity Indices at Mammoth Mountain". Proceeding International Snow Science Workshop, ISSW'94, P.O. Box 49, Snowbird, Utah, USA, 285-294, 1994.
- [13] Schweizer, J., Fohn, P.M.B., "Two expert systems to forecast the avalanche hazard for a given region". International Snow Science Workshop, Snowbird, Utah, USA, 1994.
- [14] Pant, L.M., and Ganju, A., "Fuzzy rule-based system for prediction of direct action avalanches, Current Science, Vol. 87, No.1, pp 99-104, 2004.
- [15] Gassner, M., and Brabec, B., "Nearest neighbour models for local and regional avalanche forecasting", Natural Hazards and Earth System Sciences, 2: pp 247-253, 2002.
- [16] Singh, D., Ganju, A. and Singh, A., "Weather prediction using nearest-neighbour model", Current Science, Vol. 88, No. 8, pp 1283-1289, 2005.
- [17] Singh A, and Ganju A., "A supplement to nearest-neighbour method of avalanche forecasting", Cold Region Science and Technology (39), pp.105-113, 2004.
- [18] Han, J, Kamber, M., "Data Mining: Concepts and Techniques", Morgan Kaufmann Publisher, 2001.
- [19] Yu, L., Wang, S., Lai, K.K., "An Integrated Data Preparation Scheme for Neural Network Data Analysis", IEEE Transactions on Knowledge and Data Engineering, Vol. 18, No. 2, February 2006.