International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 09 Issue 12 ǁ 2021 ǁ PP. 11-19

Snow Avalanche Prediction using Artificial Intelligence and comparative review of Algorithms

Sumit KR Sharma
Department of Computer Science Engineering
Defence Institute of Advance Technology (Deemed University)

Abstract— *Snow avalanche as a natural disaster severely affects socio-economic and geomorphic processes through damaging ecosystems, vegetation, landscape, infrastructures, transportation networks, and human life. In this study machine learning algorithms were used to predict the occurrence of avalanche using datasets containing metrological data. Classification algorithms are applied to many different problems. While these algorithms do not guarantee a definite end result, they give a solution that is close to the best result in a reasonable time. In this study, it is aimed to compare the 14 supervised machine learning algorithm under the same conditions and 2 different datasets of snow avalanche meteorological data .Thus , In order to get the best predictive model, the different classification algorithms were evaluated based on the performance metrics; accuracy, precision, recall, F1 score and AUC. For this purpose, AdaBoost Classifier, Hybrid Ensemble Model, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANNs), Stochastic Gradient Boosting algorithm, Principal Component Analysis (PCA), Naïve Bayes Classifier, Stochastic Gradient Descent algorithm, Voting Ensemble, Logistic Regression, Random Forest Classification, Decision Tree Classification and Support Vector Machine (SVM) with multiple kernels have been determined. When the result obtained from the comparison is evaluated, the best accuracy of the Snow Avalanche for both datasets is done with ANNs; the best results are obtained with ANNs for the small dataset and Random Forest Classifier for the large dataset. The obtained results using ANNs demonstrated that the proposed method outperformed in classification.*

Keywords—Snow Avalanche, prediction, supervised machine learning, classification, accuracy, algorithm

---

Date of Submission: 20-11-2021 Date of acceptance: 05-12-2021

I. INTRODUCTION

Avalanche (correspondingly known as snow slide) is basically a consequence that befalls once a cohesive snow Slab be positioned onto fragile layers of snow fracture as well as it glides downhill slope that is steeper. Triggering of the avalanches basically starts from the zones of mechanical failures in snow packs (slab avalanches) once the pressure of snow surpasses its strength however occasionally it only happens to be a gradual widening (snow avalanches that is loose). Once it has initiated, the avalanche generally accelerates promptly besides growing in volume and mass by means of entraining further snow in it. If, avalanches travel sufficiently wilder, a range of snow might combine together with air creating a dust of snow avalanche, which basically is a kind of current of gravity. [1] Avalanches generally cover three major topographies that is the zone from where it has started , the track of avalanche, and the zone of run out. Lunching of avalanches takes from the starting zone that is frequently the one of the utmost unstable portion of the slope, generally happens to be on height on mountains. As soon as the avalanches begins to slide down, it vigorously endures down the track of avalanche, downhill is the natural path which is followed by it. Later to an avalanche, huge clearances or else misplaced tree's chutes provide the evidences of the trajectory of avalanche. Finally, the avalanche approaches to stop in the run outzone, which is present at the bottommost of the slope, where the piling up of debris along with snow takes place [2]. Rocks slides or else debris of rocks also behave in a comparable manner as the snow behaves, and so it is also denoted as avalanche. Gravity may only be responsible for the burden on the snowpack, which in case of non-fulfilment might either cause weakening of the snow packs or else increase burden because of precipitation. Initiation of avalanches caused by this very method, are called spontaneous avalanches. Avalanches might as well be generated by various other charging circumstances including biological or human linked events. Activities of the seismic waves might as well generate the failures in the snow packs as well as avalanches [1].

Track of avalanche is the trails followed by an avalanche going downhill. Huge upright swath of trees misplaced from slopes or the clearings that are chute-like are seldom considered as the sign that a huge avalanche has frequently crossed over that place, making their peculiar paths. There might also be presence of big pile-ups of debris and snow found in the vicinity of bottom of slope, indicative of avalanche run through. The zone of run out is the one wherever the debris and snow have conclusively stopped. Correspondingly, this as well is considered the site of the zone of deposition, where there is the highest heaps of debris and snow. Various elements might influence the probability of avalanches, involving temperature, weather and steepness of slope, alignment of slope (whether the slope is north facing or south facing), terrain, direction of wind, usual conditions of snowpack and vegetation there. Various blending of such factors or elements can generate extreme, moderate or low avalanche situations. Several such situations, namely snowpack and temperature, might alter on an hourly or daily basis [3]. This problem is addressed by using machine learning algorithms particularly the supervised learning to help in predicting occurrence of snow avalanche in mountains.

Machine learning enables to analyse the huge amount of data and extract the patterns which can be used for different applications. The evaluation of the historical data is termed in machine learning as training of the Computational model [2]. Based on the experience collected from the historical records the future trends and upcoming events are predicted or approximated. Therefore, the data mining techniques supports the classification and prediction based on supervised learning concept for analysis of previous data. Basically the data mining techniques offered the analysis of the patterns of data and utilize them to develop classification, prediction and pattern recognition data models [3]. The supervised learning functions in two major modules training and testing. Training process evaluates the data pattern and during the testing the algorithm recognize the similar pattern data. The main advantage to use the supervised technique is their performance and accurate outcomes as compared to unsupervised approaches of learning [4]. Choosing the best model for a particular type of data mining is challenging that is why it is necessary to perform different techniques in order to choose the one that gives the best result. Therefore, the different algorithms should be evaluated based on their performance to delineate how good the predictive model is.

In this study, supervised machine learning technique was used to analyse the snow avalanche datasets that were extracted from the internet. The knowledge gained from analysis of the datasets is very useful and it will support and assist the local government unit to determine the occurrence of snow avalanche in mountain regions. This will also help them in designing effective early warning and casualty evacuation policies and programs. Different supervised machine learning technique will be evaluated and compared like Naïve Bayes, Random Forest, Decision Tree, Logistic Regression , LDA ,PCA ,SVM ,ANNs and Adaptive Boosting. To get the best predictive model, the different classification algorithms will be evaluated based on the accuracy, precision, recall, F1 score and AUC. Google Colaboratory was the coding platform used to process the different classification algorithms and to analyse the dataset.

II. LITERATURE REVIEW

A. Classification models used for different studies for classification and prediction.

The techniques that were reviewed were Naïve Bayes, Random Forest, Decision Tree, Logistic Regression, LDA, PCA, SVM, ANNs and Adaptive Boosting.

The Naive Bayes was used for crime prediction and classification in San Francisco. In the Naive Bayes, the gaussian, bernoulli, and multinomial techniques were tested. Validation and cross validation were used to test the result of each technique. The experimental results showed that higher classification accuracy can be obtained using the multinomial Naive Bayes cross validation [5]

Another study used classification method to search alternative design to simulate energy use by a building prior to the erection of the building. The classifiers used were Naïve Bayes, Decision Tree, and k-Nearest Neighbour. The study showed that Decision Tree has the fastest classification time followed by Naïve Bayes and k-Nearest Neighbour. Based on precision, recall, f-measure, accuracy, and AUC, the performance of Naïve Bayes was the best. It outperformed Decision Tree and k-Nearest Neighbour on all parameters except precision [6].

The tree-structured method has been used in the study to classify different outcomes of avalanche and it gives a fairly accurate result. The benefit of the classification tree model is simpler and efficient. Instead of using a multiple regression equation they apply the classification tree incorporating survey weights for the prediction [7].

Another research implement and compare the results of several machine learning classification algorithms such as Random Forest, Support Vector Machine, and Logistic Regression to identify poverty for different block groups in the United States. Random Forests outperform Logistic Regression and Support Vector Machines. It is observed that random forest classifier predicts well and has better performance metrics. It is also observed that random forests are relatively faster when compared to other algorithms. Random Forests are scalable and can be used with huge dataset and dimensions/features, since random forests can easily be run in a distributed environment and there is no data dependency between different trees that are formed as part of the random forest ensemble [8].

Another research implement and compare the results of performance of classification based on PCA, Linear SVM, and Multi-kernel SVM on automatic or semiautomatic measurement of various a priori brain Region of Interest (ROI) to compare and discriminate between healthy controls (HC) and Alzheimer Disease

(AD) patients. The proposed diagnosis method results yield up to about 84% stratification accuracy with Multi-kernel SVM along with high sensitivity and specificity above 85%.[15]

B. Classification of Supervised Learning Algorithms

According to [18], the supervised machine learning algorithms which deals more with classification includes the following: Linear Classifiers, Logistic Regression, Naïve Bayes Classifier, Perceptron, Support Vector Machine; Quadratic Classifiers, K-Means Clustering, Boosting, Decision Tree, Random Forest (RF); Neural networks, Bayesian Networks and so on.

1) Linear Classifiers:

Linear models for classification separate input vectors into classes using linear (hyperplane) decision boundaries [19]. The goal of classification in linear classifiers in machine learning is to group items that have similar feature values, into groups. [20] stated that a linear classifier achieves this goal by making a classification decision based on the value of the linear combination of the features. A linear classifier is often used in situations where the speed of classification is an issue, since it is rated the fastest classifier [18].Also, linear classifiers often work very well when the number of dimensions is large, as in document classification, where each element is typically the number of counts of a word in a document. The rate of convergence among data set variables however depends on the margin. Roughly speaking, the margin quantifies how linearly separable a dataset is, and hence how easy it is to solve a given classification problem [21].

2) Logistic regression:

This is a classification function that uses class for building and uses a single multinomial logistic regression model with a single estimator. Logistic regression usually states where the boundary between the classes exists, also states the class probabilities depend on distance from the boundary, in a specific approach. This moves towards the extremes (0 and 1) more rapidly when data set is larger. These statements about probabilities which make logistic regression more than just a classifier. It makes stronger, more detailed predictions, and can be fit in a different way; but those strong predictions could be wrong. Logistic regression is an approach to prediction, like Ordinary Least Squares (OLS) regression. However, with logistic regression, prediction results in a dichotomous outcome [22]. Logistic regression is one of the most commonly used tools for applied statistics and discrete data analysis. Logistic regression is linear interpolation[23].

3) Naive Bayesian (NB) Networks:

These are very simple Bayesian networks which are composed of directed acyclic graphs with only one parent (representing the unobserved node) and several children (corresponding to observed nodes) with a strong assumption of independence among child nodes in the context of their parent [24].Thus, the independence model (Naive Bayes) is based on estimating [25]. Bayes classifiers are usually less accurate that other more sophisticated learning algorithms (such as ANNs).However, [26] performed a large-scale comparison of the naive Bayes classifier with state-of-the-art algorithms for decision tree induction, instance-based learning, and rule induction on standard benchmark datasets, and found it to be sometimes superior to the other learning schemes, even on datasets with substantial feature dependencies. Bayes classifier has attribute independence problem which was addressed with Averaged One-Dependence Estimators [27].

4) Multi-layer Perceptron:

This is a classifier in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training [28].Other well-known algorithms are based on the notion of perceptron . Perceptron algorithm is used for learning from a batch of training instances by running the algorithm repeatedly through the training set until it finds a prediction vector which is correct on all of the training set. This prediction rule is then used for predicting the labels on the test set [28].

5) Support Vector Machines (SVMs):

These are the most recent supervised machine learning technique [29].Support Vector Machine (SVM) models are closely related to classical multilayer perceptron neural networks .SVMs revolve around the notion of a ―margin‖—either side of a hyperplane that separates two data classes. Maximizing the margin and thereby creating the largest possible distance between the separating hyperplane and the instances on either side of it has been proven to reduce an upper bound on the expected generalisation error [28].

6) K-means:

According to [16] and [30] K means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. K-Means algorithm is be employed when labelled data is not available [31].General method of converting rough rules of thumb into highly accurate prediction rule. Given ―weak‖ learning algorithm that can consistently find classifiers (―rules of thumb‖) at least slightly better than random, say, accuracy _ 55%, with sufficient data, a boosting algorithm can provably construct single classifier with very high accuracy, say, 99% [32].

7) Decision Trees:

Decision Trees (DT) are trees that classify instances by sorting them based on feature values. Each node in a decision tree represents a feature in an instance to be classified, and each branch represents a value that the node can assume. Instances are classified starting at the root node and sorted based on their feature values [28].Decision tree learning, used in data mining and machine learning, uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. More descriptive names for such tree models are classification trees or regression trees [33].Decision tree classifiers usually employ post-pruning techniques that evaluate the performance of decision trees, as they are pruned by using a validation set. Any node can be removed and assigned the most common class of the training instances that are sorted to it [28].

8) Neural Networks:

[32] Opined Neural Networks (NN) that can actually perform a number of regression and/or classification tasks at once, although commonly each network performs only one. In the vast majority of cases, therefore, the network will have a single output variable, although in the case of many-state classification problems, this may correspond to a number of output units (the post-processing stage takes care of the mapping from output units to output variables).Artificial Neural Network (ANN) depends upon three fundamental aspects, input and activation functions of the unit, network architecture and the weight of each input connection. Given that the first two aspects are fixed, the behaviour of the ANN is defined by the current values of the weights. The weights of the net to be trained are initially set to random values, and then instances of the training set are repeatedly exposed to the net. The values for the input of an instance are placed on the input units and the output of the net is compared with the desired output for this instance. Then, all the weights in the net are adjusted slightly in the direction that would bring the output values of the net closer to the values for the desired output. There are several algorithms with which a network can be trained [17].

9) Bayesian Network:

A Bayesian Network (BN) is a graphical model for probability relationships among a set of variables (features). Bayesian networks are the most well-known representative of statistical learning algorithms [28].The most interesting feature of BNs, compared to decision trees or neural networks, is most certainly the possibility of taking into account prior information about a given problem, in terms of structural relationships among its features [28].A problem of BN classifiers is that they are not suitable for datasets with many features [34].This prior expertise, or domain knowledge, about the structure of a Bayesian network can take the following forms:
1. Declaring that a node is a root node, i.e., it has no parents.
2. Declaring that a node is a leaf node, i.e., it has no children.
3. Declaring that a node is a direct cause or direct effect of another node.
4. Declaring that a node is not directly connected to another node.
5. Declaring that two nodes are independent, given a condition-set.
6. Providing partial nodes ordering, that is, declare that a node appears earlier than another node in the ordering.
7. Providing a complete node ordering.

C. Features of Machine Learning Algorithms

Supervised machine learning techniques are applicable in numerous domains. A number of Machine Learning (ML) application oriented papers can be found in [21], [35]. Generally, SVMs and neural networks tend to perform much better when dealing with multi-dimensions and continuous features. On the other hand, logic-based systems tend to perform better when dealing with discrete/categorical features. For neural network models and SVMs, a large sample size is required in order to achieve its maximum prediction accuracy whereas NB may need a relatively small dataset. There is general agreement that k-NN is very sensitive to irrelevant features: this characteristic can be explained by the way the algorithm works. Moreover, the presence of irrelevant features can make neural network training very inefficient, even impractical. Most decision tree algorithms cannot perform well with problems that require diagonal partitioning. The division of the instance space is orthogonal to the axis of one variable and parallel to all other axes. Therefore, the resulting regions after partitioning are all hyper rectangles. The ANNs and the SVMs perform well when multi-collinearity is present and a nonlinear relationship exists between the input and output features. Naive Bayes (NB) requires little storage space during both the training and classification stages: the strict minimum is the memory needed to store the prior and conditional probabilities. The basic kNN algorithm uses a great deal of storage space for the training phase, and its execution space is at least as big as its training space. On the contrary, for all non-lazy learners, execution space is usually much smaller than training space, since the resulting classifier is usually a highly condensed summary of the data. Moreover, Naive Bayes and the kNN can be easily used as incremental learners whereas rule algorithms cannot. Naive Bayes is naturally robust to missing values since these are simply ignored in computing probabilities and hence have no impact on the final decision. On the contrary, kNN and neural networks require complete records to do their work. Finally, Decision Trees and NB generally have different operational profiles, when one is very accurate the other is not and vice versa. On the contrary, decision trees and rule classifiers have a similar operational profile. SVM and ANN have also a similar operational profile. No single learning algorithm can uniformly outperform other algorithms over all datasets. Different data sets with different kind of variables and the number of instances determine the type of algorithm that will perform well. There is no single learning algorithm that will outperform other algorithms based on all data sets according to no free lunch theorem. [36] Table 1 presents the comparative analysis of various learning algorithms.

III. METHODS

The methodological approach of this study is composed of: the collection of data sets about snow avalanche, identification of metrological factors to predict occurrence of avalanche, formulation of the predictive model using the supervised machine learning algorithms, andthe performance evaluation metrics applied to determine the best predictive model.

A. Data Collection

The snow avalanche data was collected from the internet. Two datasets were collected with multiple features and different number of data points. The dataset with 204 data points and 5 features, 4216 data points and 3 features are referred as small and large datasets respectively in this study. Both the datasets are having following metrological and geographical features: temperature, snowfall, sunlight, forest density, slope and intermittent snow.

B. Data Pre-processing

The avalanche datasets were rich in terms of variables and number of data points therefore it was pre-processed to find the co-relation among the attributes which enable better understanding of their impact on prediction. This process also includes cleaning the data set that involves dealing with missing values, outliers and inconsistent values. Missing values were replaced by most used value in the data while the inconsistent values were fixed.

The values were normalized for converting it into three classes to achieve algorithms to perform classification. The data was subdivided into two, 80% of the data was used as the training data while the remaining 20% was used as the test data. The division of data is necessary in order to get the best predictive model in predicting the three classes of occurrence of snow avalanche. The correlation of variables with prediction is summarized in table 1.

Table 1: correlation of variables with prediction variable

Dataset	Variables	Correlation Value
1.Small Dataset (5 Features And 204 Data Points)	Slope	0.487693
	Snow Density	0.454154
	Air Temperature	-0.200510
	Wind	-0.014613
1.Large Dataset (4 Features And 4216 Data Points)	Temperature	-0.576589
	Relative Humidity	0.113542
	Total Snow Depth	0.661571

C. Data Analysis

Data analysis involves manipulating, transforming, and visualizing data in order to infer meaningful insights from the results. Mathematical tools like bar graphs, heat maps, histograms and correlation matrix were used to visualize both the datasets;These tools gave an insight about features which were more relevant in predicting class.

Until the final decades of the 20th century, human analysts were irreplaceable when it came to finding patterns in data. Today, they're still essential when it comes to feeding the right kind of data to learning algorithms and inferring meaning from algorithmic output, but machines can and do perform much of the analytical work itself.

Data analysis allowed to analyse data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational database [22].

The googlecolaboratory was used for coding for this study to predict the classes for avalanche prediction. The mathematical tools from different libraries of python language were used to determine the classification and to determine the performance of the fourteen classification algorithms like Naïve Bayes, Decision Tree, ANN, Logistic Regression and PCA .

D. Performance Evaluation

To get the best predictive model in prediction, the different classification algorithms were evaluated based on the following:

A. Accuracy that measures the performance of each model that gives percentage of features that are predicted correctly among total number of features,

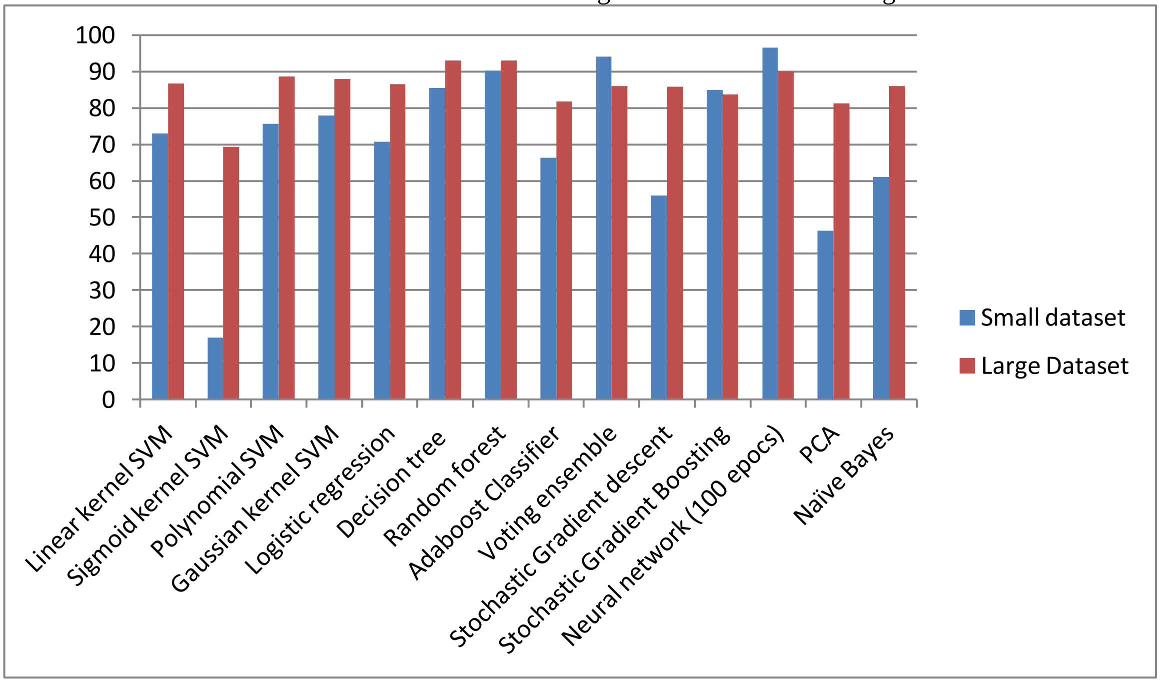
B. Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

C. Recall that gives number of positive features classified correctly by the model,

D. F1 Score is a harmonic mean of precision and recall for balancing out both has been taken as a measure of performance.

E. AUC provides an aggregate measure of performance across all possible classification thresholds. One way of interpreting AUC is as the probability that the model ranks a random positive example more highly than a random negative example.

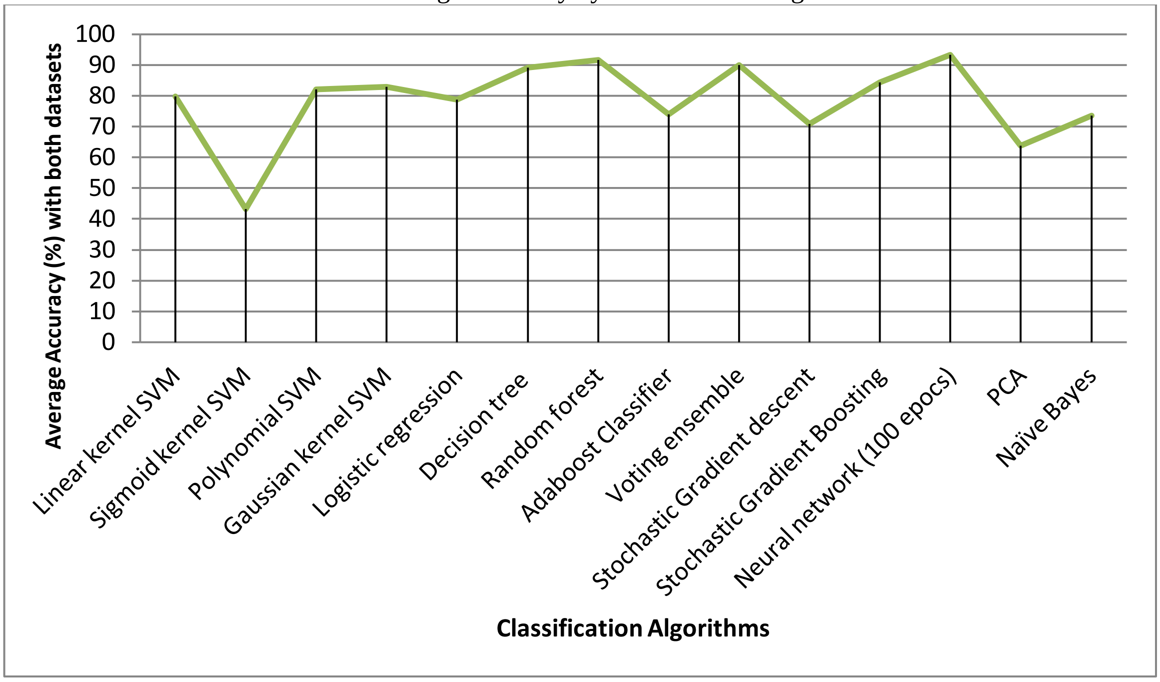
Chart 1: Performance of Classification algorithms on Small and Large Datasets



IV. RESULTS AND DISCUSSION

Chart 1 shows the performance of the fourteen classification algorithms used in this study with the comparator as accuracy score.WhileChart 2 shows the comparative graph of the all algorithms based on their performance metrics.

Chart 2: Average Accuracy by Classification Algorithms



The algorithms AdaBoost Classifier, Hybrid Ensemble Model, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANNs), Stochastic Gradient Boosting algorithm, Principal Component Analysis (PCA), Naïve Bayes Classifier, Stochastic Gradient Descent algorithm, Voting Ensemble, Logistic Regression, Random Forest Classification, Decision Tree Classification and Support Vector Machine (SVM) with multiple kernels were used in this study to predict the occurrence of snow avalanche. These fourteen algorithms were evaluated based on their performance metrics to get the best predictive model. It can be seen in Table 2 that based on the accuracy of the fourteen algorithms the ANNs out performs the other algorithms with an accuracy rate of 96.5 and 90% for small and large dataset respectively , while the Sigmoid kernel SVM has the lowest accuracy rate of 17 and 69.3% for small and large dataset respectively.

The performance of classification algorithms was compared primarily based on accuracy. The Sigmoid Kernel SVM, Stochastic Gradient Descent, PCA and Naïve Bayes algorithm displayed inconsistent performance for the small dataset. From this performance metric, it prevails that ANN has the highest accuracy measure of 96.5% for small dataset while the Decision Tree and Random Forest classifier got the highest accuracy measure rate of 93% for large dataset. Precision, Recall and F1 score were also calculated for each classifier and both datasets, Accuracy score was considered primary comparator because of critical nature of application.

Table 2: Classification Algorithms Performance

ClassificationAlgorithm	**Accuracy in %**	
	Small dataset	Large Dataset
Linear kernel SVM	73	86.72
Sigmoid kernel SVM	17	69.3
Polynomial SVM	75.6	88.6
Gaussian kernel SVM	78.0	88.0
Logistic regression	70.7	86.6
Decision tree	85.4	**93.0**
Random forest	90.2	**93.1**
Adaboost Classifier	66.3	81.8
Voting ensemble	94	86
Stochastic Gradient descent	56.0	85.8
Stochastic Gradient Boosting	85	83.8
Neural network (100 epocs)	**96.5**	90
PCA	46.3	81.2
Naïve Bayes	61	86

Furthermore, Chart 2 shows the comparative graph of the fourteen algorithms and it shows that Artificial Neural Networks predict well and has the best performance metrics compare to the other classification algorithms used in this study.

The result of the study could help the local government unit to determine the occurrence of snow avalanche in mountain regions. This will also help them in designing effective early warning and casualty evacuation policies and programs. The result can also help the policy makers to easily identified appropriate casualty reduction policies and programs to be rendered in the mountain area.

V. CONCLUSION

Choosing the right algorithm for a particular type of classification task is not easy. The best way is to perform a validation of the performances of the different algorithms to choose the best one that gives an appropriate result. This study performed a comparative analysis of the fourteen classification algorithms; AdaBoost Classifier, Hybrid Ensemble Model, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANNs), Stochastic Gradient Boosting algorithm, Principal Component Analysis (PCA), Naïve Bayes Classifier, Stochastic Gradient Descent algorithm, Voting Ensemble, Logistic Regression, Random Forest Classification, Decision Tree Classification and Support Vector Machine (SVM) with multiple kernels using the snow avalanche data setstaken from the internet . In order to get the best predictive model, the different classification algorithms were evaluated based on the performance metrics; accuracy, precision, recall, F1 score and AUC. The Google Colaboratory was used to process thedifferent classification algorithms and to analyse the dataset. It can be concluded that ANNs are efficient for predicting occurrence of snow avalanche because it outperformed all the other algorithms in all the performance metrics used in this study. Furthermore, the error rate of the said algorithm is significant because it is 0.09 only. Therefore, choosing the best classification algorithm for classification task will give a better result.

ACKNOWLEDGMENT

The researcher wishes to express his genuine appreciation to all those who are, in one way or another, instrumental to the successful completion of this work.
Dr.UpasnaSingh , Assistant Professor and member DRMC, Department of CSE , DIAT(DU), for her guidance on Machine Learning Subjects.
Dr.ManishaNene , Head and DRMC Chairman, Department of CSE , DIAT(DU), for her guidance on Intelligent algorithms.

REFERENCES

[1]. Raza, Muhammad Qamar, and Abbas Khosravi. "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings." *Renewable and Sustainable Energy Reviews* 50 (2015): 1352-1372.

[2]. Metaxiotis, K., et al. "Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher." *Energy conversion and Management* 44.9 (2003): 1525-1534.

[3]. Choubin, Bahram, et al. "Snow avalanche hazard prediction using machine learning methods." *Journal of Hydrology* 577 (2019): 123929.

[4]. Chouhan, S. S., &Khatri, R. (2016). Data Mining based Technique for Natural Event Prediction and Disaster Management. International Journal of Computer Applications, 139(14).

[5]. Abdulrahman, N., &Abedalkhader, W. KNNCLASSIFIER AND NAÏVE BAYSE CLASSIFIER FOR CRIME PREDICTION IN SAN FRANCISCO CONTEXT.

[6]. Jeyarani, D. S., Anushya, G., &Pethalakshmi, A. (2013). A comparative study of decision tree and Naive Bayesian classifiers on medical datasets.age, 30, 31-40.

[7]. Bilton, P., Jones, G., Ganesh, S., & Haslett, S. (2017). Classification trees for poverty mapping. Computational Statistics & Data Analysis, 115, 53-66.

[8]. Korivi, K. (2016). Identifying poverty-driven need by augmenting census and community survey data.

[9]. Tolan, G. M., &Soliman, O. S. (2015). An Experimental Study of Classification Algorithms for Terrorism Prediction. International Journal of Knowledge Engineering-IACSIT, 1(2), 107-112.

[10]. Ashari, A., Paryudi, I., &Tjoa, A. M. (2013). Performance comparison between Naïve Bayes, decision tree and k-nearest neighbor in searching alternative design in an energy simulation tool.International Journal of Advanced Computer Science and Applications (IJACSA), 4(11).

[11]. Gohar, F., Butt, W. H., &Qamar, U. (2014). Terrorist Group Prediction Using Data Classification.In Work.MultiRelational Data Min. MRDM2003 (Vol. 10, pp. 199-208).

[12]. Adhatrao, K., Gaykar, A., Dhawan, A., Jha, R., &Honrao, V. (2013). Predicting students' performance using ID3 and C4.5 classification algorithms.arXiv preprint arXiv:1310.2071.

[13]. Pandey, A. (2014). Study and Analysis of K-Means Clustering Algorithm Using Rapidminer. International Journal of Engineering Research and Applications, 4(12), 60-64.)

[14]. COE, J. (2012). Performance comparison of Naïve Bayes and J48 classification algorithms.International Journal of Applied Engineering Research, 7(11), 2012.

[15]. S. Alam, Moonsoo Kang, Jae-Young Pyun and G. Kwon, "Performance of classification based on PCA, linear SVM, and Multi-kernel SVM," 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), 2016, pp. 987-989, doi: 10.1109/ICUFN.2016.7536945.

[16]. Bishop, C. M. (1995). Neural Networks for Pattern Recognition. Clarendon Press, Oxford, England. 1995. Oxford University Press, Inc. New York, NY, USA ©1995 ISBN:0198538642 Available at: http://cs.du.edu/~mitchell/mario_books/Neural_Networks_for_Pattern_Recognition_-_Christopher_Bishop.pdf

[17]. Neocleous C. &Schizas C. (2002). Artificial Neural Network Learning: A Comparative Review. In: Vlahavas I.P., Spyropoulos C.D. (eds)Methods and Applications of Artificial Intelligence. Hellenic Conference on Artificial IntelligenceSETN 2002. Lecture Notes in Computer Science, Volume 2308. Springer, Berlin, Heidelberg, doi: 10.1007/3-540-46014-4_27 pp. 300-313. Available at: https://link.springer.com/chapter/10.1007/3-540-46014- 4_27

[18]. Taiwo, O. A. (2010). Types of Machine Learning Algorithms, New Advances in Machine Learning, Yagang Zhang (Ed.), ISBN: 978-953-307-034-6, InTech, University of Portsmouth United Kingdom. Pp 3 – 31. Available at InTech open website: http://www.intechopen.com/books/new-advances-in-machine-learning/types-of-machine-learning-algorithms

[19]. Elder, J. (n.d). Introduction to Machine Learning and Pattern Recognition. Available at LASSONDE University EECS Department York website: http://www.eecs.yorku.ca/course_archive/2011-12/F/4404-5327/lectures/01%20Introduction.pd

[20]. Timothy Jason Shepard, P. J. (1998). Decision Fusion Using a Multi-Linear Classifier.In Proceedings of the International Conference on Multisource-Multisensor Information Fusion.

[21]. Setiono R. and Loew, W. K. (2000), FERNN: An algorithm for fast extraction of rules from neural networks, Applied Intelligence.

[22]. Newsom, I. (2015). Data Analysis II: Logistic Regression. Available at: http://web.pdx.edu/~newsomj/da2/ho_logistic.pdf

[23]. Logistic Regression pp. 223 – 237. Available at: https://www.stat.cmu.edu/~cshalizi/uADA/12/lectures/ch12.pdf

[24]. Good, I.J. (1951). Probability and the Weighing of Evidence, Philosophy Volume 26, Issue 97, 1951. Published by Charles Griffin and Company,

[25]. Nilsson, N.J. (1965). Learning machines. New York: McGraw-Hill.Published in: Journal of IEEE Transactions on Information Theory Volume 12 Issue 3, 1966. doi: 10.1109/TIT.1966.1053912 pp. 407 – 407. Available at ACM digital library website: http://dl.acm.org/citation.cfm?id=2267404

[26]. Domingos, P. &Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss.Machine Learning Volume 29, pp. 103–130 Copyright © 1997 Kluwer Academic Publishers.Manufactured in The Netherlands. Available at University of Trento website: http://disi.unitn.it/~p2p/RelatedWork/Matching/domingos97optimality.pdf

[27]. Hormozi, H., Hormozi, E. &Nohooji, H. R. (2012). The Classification of the Applicable Machine Learning Methods in Robot Manipulators. International Journal of Machine Learning and Computing (IJMLC), Vol. 2, No. 5, 2012 doi: 10.7763/IJMLC.2012.V2.189pp. 560 – 563. Available at IJMLC website: http://www.ijmlc.org/papers/189-C00244-001.pdf

[28]. Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques.Informatica 31 (2007). Pp. 249 – 268. Retrieved from IJS website: http://wen.ijs.si/ojs-2.4.3/index.php/informatica/article/download/148/140.

[29]. Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. (2nd ed.). Springer Verlag. Pp. 1 – 20. Retrieved from website: https://www.andrew.cmu.edu/user/kk3n/simplicity/vapnik2000.pdf

[30]. Tapas Kanungo, D. M. (2002). A local search approximation algorithm for k-means clustering.Proceedings of the eighteenth annual symposium on Computational geometry. Barcelona, Spain: ACM Press.

[31]. Alex S.&Vishwanathan, S.V.N. (2008). Introduction to Machine Learning.Published by the press syndicate of the University of Cambridge, Cambridge, United Kingdom.Copyright ⓒ Cambridge University Press 2008. ISBN: 0-521-82583-0. Available at KTH website: https://www.kth.se/social/upload/53a14887f276540ebc81aec3/online.pdf Retrieved from website: http://alex.smola.org/drafts/thebook.pdf

[32]. Rob Schapire (n.d) Machine Learning Algorithms for Classification.

[33]. T. Hastie, R. Tibshirani, J. H. Friedman (2001) ― The elements of statistical learning,‖ Data mining, inference, and prediction, 2001, New York: Springer Verlag.

[34]. Cheng, J., Greiner, R., Kelly, J., Bell, D.& Liu, W. (2002). Learning Bayesian networks from data: An information-theory based approach. Artificial Intelligence Volume 137.Copyright © 2002.Published by Elsevier Science B.V. All rights reserved pp. 43 – 90. Available at science Direct: http://www.sciencedirect.com/science/article/pii/S0004370202001911

[35]. Witten, I. H. & Frank, E. (2005). Data Mining: Practical machine learning tools and techniques (2nd ed.), ISBN: 0-12-088407-0, Morgan Kaufmann Publishers, San Francisco, CA, U.S.A. © 2005 Elsevier Inc.Retrieved from website: ftp://93.63.40.27/pub/manuela.sbarra/Data Mining Practical Machine Learning Tools and Techniques - WEKA.pdf

[36]. Lemnaru C. (2012). Strategies for dealing with Real World Classification Problems, (Unpublished PhD thesis) Faculty of Computer Science and Automation, UniversitateaTechnica, Din Cluj-Napoca. Available at website: http://users.utcluj.ro/~cameliav/documents/TezaFinalLemnaru.pdf

[37]. Pradeep, K. R. & Naveen, N. C. (2017). A Collective Study of Machine Learning (ML)Algorithms with Big Data Analytics (BDA) for Healthcare Analytics (HcA). International Journal of Computer Trends and Technology (IJCTT) – Volume 47 Number 3, 2017. ISSN: 2231-2803, doi: 10.14445/22312803/IJCTT-V47P121, pp 149 – 155. Available from IJCTT website: http://www.ijcttjournal.org/2017/Volume47/number-3/IJCTT-V47P121.pdf