International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Sonar Target Classification Problem: Machine Learning Models

Ritwick Ghosh

UG Student at Indian Institute of Engineering Science and Technology, Shibpur, Howrah, West Bengal, India, Pin: 711103

Abstract: ***In this study various machine learning algorithms are used for a noisy binary classification problem. The dataset that was used by Gorman and Sejnowski (1988) in their study of the classification of sonar signals using a neural network of undersea targets is used in this study. The task was to train a network to discriminate between sonar signals bounced off a metal cylinder and those bounced off a roughly cylindrical rock. The data used for the network experiments were sonar returns collected from a metal cylinder and a cylindrically shaped rock positioned on a sandy ocean floor. This dataset has 60 different features and is extremely noisy in nature. Total 29machine classification algorithms are used on the dataset. Programming languages used in this study are MATLAB 2018b and Python 3.7.***

Keywords: Sonar, Machine Learning, KNN, Decision Tree, Linear Regression, Ensemble

1. Introduction

Sonaris a technique that uses sound propagation characteristics to detect objects by emitting sound pulses and detecting or measuring their return after reflection, used as a means of acoustic location by measuring of the echo characteristics of object. Machine learning is the statistical study of computer algorithms to perform a specific task depending upon patterns without explicit instructions. It is also considered as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on given data, in order to make predictions or decisions without being explicitly programmed to perform the task for any new non-classified data point. In the present study, machine learning algorithms are applied to a sonar target classification problem. The models were trained to classify sonar returns from an undersea metal cylinder and a cylindrically shaped rock, comparable in size, using various algorithms. Reported dataset[6] is taken from another paper (Gorman & Sejnowski, 1988) [5]. Finally; a comparison is drawn between trained classifiers and human listeners trained to discriminate the same two classes of sonar returns.

2. Related Works

In the paper of Gorman & Sejnowski (1988), parallel networks were trained to classify undersea rock and metal according to sonar returns. The dataset was explained and evaluated by another paper (Gorman & Sejnowski, 1987) [4]. On a sandy ocean floor the 5 ft. long meta cylinder and rock was placed and wide-band linear and frequency-modulated FM chirps (ka =55.6) (rising in frequency) impinging pulse was shot towards them from around 10 meter distance, aspect angles ranging from 90^0 to 180^0. Each pattern is a set of 60 numbers in the range of 0.0 to 1.0 and every number ranging in between 0.0 to 1.0, represents the energy within a particular frequency band, integrated over a certain period of time. There were total 208 patterns in the prepared dataset, divided into two classes- metal (111 patterns) and rock cylinder (97 patterns). They used neural networks with nearest neighbor solver and nearest neighbor classifier which had accuracy of 90.4% and 82.7%.

The following dataset is cited and used by another 53 research papers (Connectionist Bench Data Set. UCI Machine Learning Repository, till October 2019) [6]. Different types of machine learning and deep learning algorithms are tested upon the dataset and the dataset fetch impressive accuracy for different classification algorithms. In this study we used machine learning algorithms only.

Properties of the Dataset

This sonar signal based dataset is considered to be an ideal example of multiclass-noisy dataset. Along with 60 different features this dataset is noisy, deceptive and seemed random. Only computational intelligence can extract the explicit pattern out of it.

3. Methodology

Python 3.7 along with Scikit-Learn module and MATLAB 2018b along with Classification Learner is used to analyze the dataset. These are the following machine learning algorithms that are used: linear algorithms (Logistic Regression), Discriminant Analysis(Linear and Quadratic Discriminant), tree algorithms (Fine tree, Medium tree, Coarse tree, Extra tree, Decision tree), Support Vector Machines/SVM (Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, Coarse Gaussian SVM), K-Nearest Neighbor/KNN (Fine, Medium, Coarse, Cosine, Cubic,Weighted KNN), Navies Bayes/NB(GaussianNB), and Ensemble (Ada Boost, Boosted trees, Bagged trees, Subspace Discriminant, Subspace KNN, RUBoost, Random Forest,Scalable Tree Gradient Boosting-XGBoosting).

Logistic regression models binary variables using logistic function. Linear discrimination tries to find a linear or quadratic combination of features for classification. Tree algorithm tries to organize the data into tree data structure depending upon its features.SVM algorithm focuses on finding the natural cluster in between the data. KNN algorithm assign the class according to the plurality vote of its neighbors' classes.NB tries to classify applying Bayes' theorem according to naïve (strong) independence assumptions in between the features. Ensemble algorithm create a concrete set of alternative statistical models to classify targets.

For every test 5-fold validation is used to nullify the chance of model over-fitting. It means the whole dataset was divided into 5 disjoint sets and the experiment continues for 5 epochs as each set (20%) to be test data and other sets (80%) to be train data.

4. Results

Table 1: Accuracy Comparison of Different Machine Learning Models

Machine Learning Algorithm	*Accuracy (%)*
Logistic Regression	76.0
Linear Discriminant	76.4
Quadratic Discriminant	72.6
Fine Tree	70.2
Medium Tree	70.2
Coarse Tree	68.3
Extra Tree	78.38
Decision Tree	73.49
Linear SVM	75.0
Quadratic SVM	84.6
Cubic SVM	87.0
Fine Gaussian SVM	59.6
Medium Gaussian SVM	83.7
Coarse Gaussian SVM	74.5
Fine KNN	85.1
Medium KNN	73.1
Coarse KNN	70.7
Cosine KNN	76.4
Cubic KNN	71.2
Weighted KNN	79.3
Gaussian NB	64.89
Ada Boost	81.39
Boosted Trees	53.4
Bagged Trees	79.3
Subspace Discriminant	77.9
Subspace KNN	82.7
RUSBoosted Trees	69.7
Random Forest	78.82
XGBoost	88.1

5. Conclusion

In the study by Gorman and Sawatari (1987) [3] human subjects were tested to classify the same targets by hearing the reflected sound wave (FM chirp, 500 to 1100Hz); 100 returns for each subject. The best performance was 88%, but according to the study the average accuracy score was 82%. Although it the results should not be compared, it is certain from the above experiment that the machine learning models accuracy is comparable with the human hearing accuracy.

This study also represents the variety of statistical models available for processing and their comparative accuracy for this binary classification problem, considering it is a noisy and deceptive dataset.

References

[1] Bhattacharyya.C.. (2004)-*Robust Classification of noisy data using Second Order Cone Programming approach*. Dept. Computer Science and Automation, Indian Institute of Science. (Retrieved from: http://rexa.info/paper/ab6823c2bce78649401e56bfb2d46bb279535879)

[2] Campbell C. and Cristianini.N. (2006)-*Simple Learning Algorithms for Training Support Vector Machines*. Dept. of Engineering Mathematics. (Retrieved from: tp://www.svms.org/training/CaCr.pdf)

[3] [3] Gorman R. P., and Sawatari T, (1987). *Automatic sonar target recognition based on human perceptual features*. The Acoustic Society of America

[4] Gorman R. P., & Sejnowski, T. J. (1987)-*Learned classification of sonar targets using a massively-parallel network*. IEEE Trans on Acoustics Speech Signal Processing. (Retrieved from: https://www.researchgate.net/publication/3175285_Learned_Classifications_of_Sonar_Targets_Using_a_Massively_Parallel_Network)

[5] Gorman, R. P., and Sejnowski, T. J. (1988) -*Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets*. Neural Networks, Vol. 1, pp. 75-89. (Retrieved from: https://papers.cnl.salk.edu/PDFs/Analysis%20of%20Hidden%20Units%20in%20a%20Layered%20Network%20Trained%20to%20Classify%20Sonar%20Targets%201988-2996.pdf)

[6] Sejnowski T. and Gorman R.P. (1988) -*UCI Machine Learning Repository*. [https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29]. Salk Institute and the University of California at San Deigo, Allied-Signal Aerospace Technology Centre.