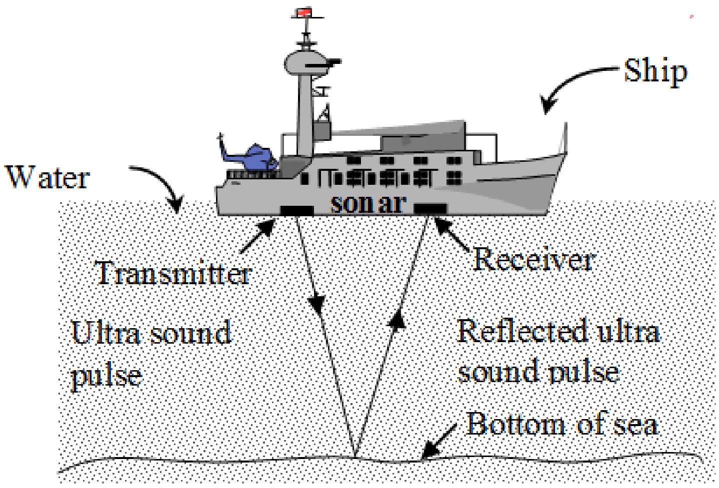
Symbiosis Skills and Professional University

Kiwale, Pune

PROJECT REPORT

On

"BINARY CLASSIFICATION OF ROCKS AND MINES USING SONAR DATA"



Submitted by

Sachin Shantaram Sabale

(Registration Number- 2001206070)

ML-Batch-III

Under The Guidance of

1) Dr. Ruby Jain 2) Mr. Sanjay Bhorekar

STUDENT DECLARATION AND ATTESTATION BY TRAINER

This is to declare that this report has been written by me. No part of the report is plagiarized from other sources. All information included from other sources has been duly acknowledged. I aver that if any part of the report is found to be plagiarized, I shall take full responsibility for it.

Signature of student

Name of student: Sachin Shantaram Sabale

Registration Number: 2001206070)

Signature of trainer

Name of trainer:

CERTIFICATE

This is to certify that the report entitled, "**Binary Classification of Rocks and Mines Using SONAR Data**" submitted by" Sachin Shantaram Sabale" to Symbiosis Skills and Professional University, Pune, Maharashtra, India, is a record of bonafide Project work carried out by him under my supervision and guidance and is worthy of consideration for the completion of certificate course in 'Machine Learning Engineer".

Signature of Trainer
Name of Trainer

Date: / / 2021

_____________________ _____________________

Supervisor Supervisor

Date:

INDEX

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions of the following people without whose help and guidance this report would not have been completed.

I acknowledge the counsel and support of our trainer Dr. Ruby Jain, Mr. Sanjay Bhorekar, Mr. Swapnil Pingale for your guidance. I acknowledge the support of our training coordinator Mr. Sumit Bhagat and placement coordinator Mr Zameer Kazi.

This acknowledgement will remain incomplete if I fail to express our deep sense of obligation to my parents and God for their consistent blessings and encouragement.

INTRODUCTION AND AIM

BINARY CLASSIFICATION OF ROCKS AND MINES USING SONAR DATA

The task is to discriminate between sonar signals bounced off a metal cylinder and those bounced off a roughly cylindrical rock. Or simply predict metal or rock objects from sonar return data.

The data set was contributed to the benchmark collection by Terry Sejnowski, now at the Salk Institute and the University of California at San Deigo. The data set was developed in collaboration with R. Paul Gorman of Allied-Signal Aerospace Technology Center.

Downloaded from: https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+(Sonar,+Mines+vs.+Rocks)

OBJECTIVES

The task is to discriminate between sonar signals bounced off a metal cylinder and those bounced off a roughly cylindrical rock. Or simply predict metal or rock objects from sonar return data.

- To find accuracy on training data
- To find accuracy on testing data
- To make predictive system
- Feature Selection to increase the performance of the model.
- Find given collection of data is of mines or rocks

DESCRIPTION OF DATASET AND METHOD

The file "sonar.mines" contains patterns obtained by bouncing sonar signals off a metal cylinder at various angles and under various conditions. The file "sonar.rocks" contains 97 patterns obtained from rocks under similar conditions. The transmitted sonar signal is a frequency-modulated chirp, rising in frequency. The data set contains signals obtained from a variety of different aspect angles, spanning 90 degrees for the cylinder and 180 degrees for the rock.

Each pattern is a set of 60 numbers in the range 0.0 to 1.0. Each number represents the energy within a particular frequency band, integrated over a certain period of time. The integration aperture for higher frequencies occurs later in time, since these frequencies are transmitted later during the chirp.

The label associated with each record contains the letter "R" if the object is a rock and "M" if it is a mine (metal cylinder). The numbers in the labels are in increasing order of aspect angle, but they do not encode the angle directly

WHAT IS "SONAR" SYSTEM

SONAR (**sound navigation and ranging**) is a technique that uses sound propagation (usually underwater, as in submarine navigation) to navigate, communicate with or detect objects on or under the surface of the water, such as other vessels.

A sonar device sends pulses of sound waves down through the water. When these pulses hit objects like fish, vegetation or the bottom, they are reflected back to the surface. The sonar device measures how long it takes for the sound wave to travel down hit an object and then bounce back up.

Today, SONAR has many uses in the maritime world, from mapping the seafloor to exploring shipwrecks. SONAR is short for Sound Navigation and Ranging. ... It was designed to detect icebergs underwater to help ships navigate around them. This detection system became more important after the sinking of the.

Sonar uses sound waves to 'see' in the water.

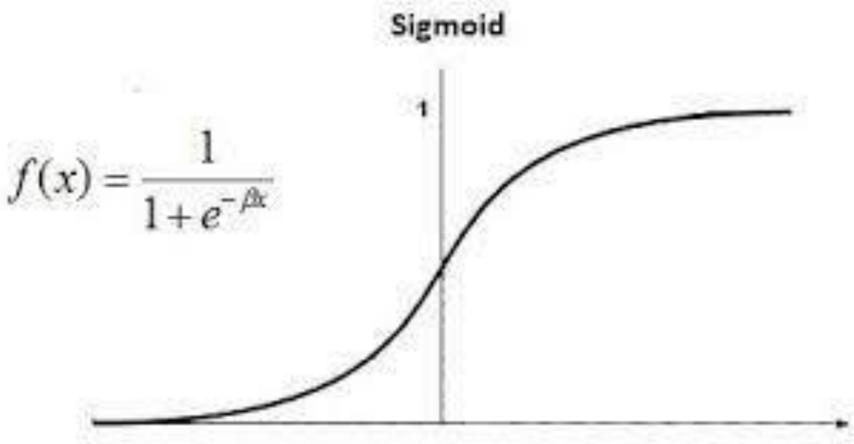
There are two types of sonar—active and passive.

CLASSIFICATION IN MACHINE LEARNING

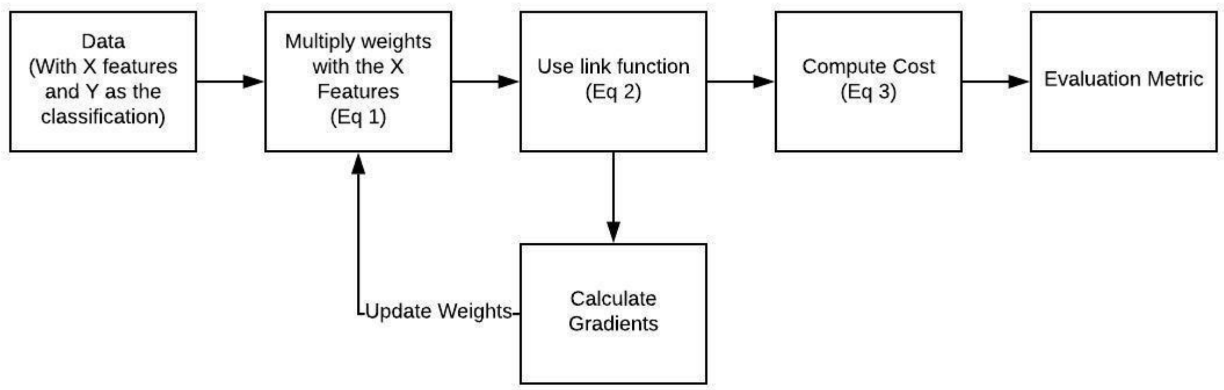
Classification and Regression are the parts of Supervise machine learning. For the binary classification Rocks and Mines there are some algorithms like Logistic Regression, Support Vector Machine, Naïve Bayes, etc. ….

Here, logistics regression is used……

LOGISTIC REGRESSION is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression).



Logistic regression Flow



Logistic Regression process

Given a data(X,Y), X being a matrix of values with m examples and n features and Y being a vector with m examples. The objective is to train the model to predict which class the future values belong to. Primarily, we create a weight matrix with random initialization. Then we multiply it by features.

$$a = w_0 + w_1x_1 + w_2x_2 + ...w_nx_n$$

Eq 1.

We then pass the output obtained from Eq 1. to a link function.

$$y\hat{i} = 1/(1 + e^{-a})$$

Link Function

This is followed by calculating the cost for that iteration whose formula is

$$cost(w) = (-1/m)\sum_{i=1}^{i=m} y_i log(\hat{y_i}) + (1 - y_i) log(1 - \hat{y_i}))$$

Cost function

The derivative of this cost is calculated following which the weights are updated.

$$dw_j = \sum_{i=1}^{i=n} (\hat{y} - y) x_j^i$$

Gradient

$$w_i = w_j - (\alpha * dw_j)$$

Update

DATA ANALYSIS, TRAINING AND TESTING

SEPARATING DATA AND LABELS..

```
# separating data and Labels
X = sonar_data.drop(columns=60, axis=1)
Y = sonar_data[60]
```

Training and Test data

In [12]:
```
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size = 0.1, stratify=Y, random_state=1)
```

In [13]:
```
print(X.shape, X_train.shape, X_test.shape)
```

```
(208, 60) (187, 60) (21, 60)
```

In [14]:
```
print(X_train)
```

```
          0       1       2       3       4       5       6       7       8   \
115  0.0414  0.0436  0.0447  0.0844  0.0419  0.1215  0.2002  0.1516  0.0818
38   0.0123  0.0022  0.0196  0.0206  0.0180  0.0492  0.0033  0.0398  0.0791
56   0.0152  0.0102  0.0113  0.0263  0.0097  0.0391  0.0857  0.0915  0.0949
123  0.0270  0.0163  0.0341  0.0247  0.0822  0.1256  0.1323  0.1584  0.2017
18   0.0270  0.0092  0.0145  0.0278  0.0412  0.0757  0.1026  0.1138  0.0794
..      ...     ...     ...     ...     ...     ...     ...     ...     ...
140  0.0412  0.1135  0.0518  0.0232  0.0646  0.1124  0.1787  0.2407  0.2682
5    0.0286  0.0453  0.0277  0.0174  0.0384  0.0990  0.1201  0.1833  0.2105
154  0.0117  0.0069  0.0279  0.0583  0.0915  0.1267  0.1577  0.1927  0.2361
131  0.1150  0.1163  0.0866  0.0358  0.0232  0.1267  0.2417  0.2661  0.4346
203  0.0187  0.0346  0.0168  0.0177  0.0393  0.1630  0.2028  0.1694  0.2328

          9   ...      50      51      52      53      54      55      56  \
115  0.1975  ...  0.0222  0.0045  0.0136  0.0113  0.0053  0.0165  0.0141
38   0.0475  ...  0.0149  0.0125  0.0134  0.0026  0.0038  0.0018  0.0113
```

```
print(Y_train)
```

```
115    M
38     R
56     R
123    M
18     R
      ..
140    M
5      R
154    M
131    M
203    M
Name: 60, Length: 187, dtype: object
```

MODEL TRAINING --> LOGISTIC REGRESSION

MODEL FITTING-

```
model = LogisticRegression()
```

```
#training the Logistic Regression model with training data
model.fit(X_train, Y_train)
```

```
LogisticRegression()
```

ACCURACY OF TRAINING AND TESTING DATA-

```
#accuracy on training data
X_train_prediction = model.predict(X_train)
training_data_accuracy = accuracy_score(X_train_prediction, Y_train)
```

```
print('Accuracy on training data : ', training_data_accuracy)
```

```
Accuracy on training data :  0.8342245989304813
```

```
#accuracy on test data
X_test_prediction = model.predict(X_test)
test_data_accuracy = accuracy_score(X_test_prediction, Y_test)
```

```
print('Accuracy on test data : ', test_data_accuracy)
```

```
Accuracy on test data :  0.7619047619047619
```

MAKING OF PREDICTIVE SYSTEM

```
input_data = (0.0307,0.0523,0.0653,0.0521,0.0611,0.0577,0.0665,0.0664,0.1460,0.2792,0.3877,0.4992,0.4981,0.4972,0.5607,0.73

# changing the input_data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the np array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

prediction = model.predict(input_data_reshaped)
print(prediction)

if (prediction[0]=='R'):
  print('The object is a Rock')
else:
  print('The object is a Mine')
```

```
['M']
The object is a mine
```

CONCLUSION

- ❖ Logistic Regression model successfully fitted.

- ❖ Accuracy on training data : 0.8342245989304813

- ❖ Accuracy on test data : 0.7619047619047619

- ❖ Predictive system classifies the given collection of SONAR data as a Mine.

FUTURE SCOPE

This work can be further extended to find the best classification algorithm using deep learning approaches.

Using deferent algorithms of Machine learning or Deep learning, classification of more than two object.

Used for training simulators and to create more productive underwater surveillance.

REFERENCES

https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+(Sonar,+Mines+vs.+Rocks)

https://www.kaggle.com/mattcarter865/mines-vs-rocks

https://web.stanford.edu/~jurafsky/slp3/5.pdf

https://morioh.com/p/4429fa67611e