

Welcome to the newly launched Education Spotlight page!

[VIEW LISTINGS](#)Join us in
ORLANDO
November 5–10New ML, AI, and Hands-On Python Courses
In-depth training, expert instructors
MLOps: Best Practices for Delivering
and Managing Machine LearningSave 30%
w/code **DSC30****Data Science Central**
A COMMUNITY FOR AI PRACTITIONERS

Home » Uncategorized

Credit Risk Prediction Using Artificial Neural Network Algorithm

**ShrutiGoyal** | March 14, 2018 at 7:30 am

1 Introduction

Credit risk or credit default indicates the probability of non-repayment of bank financial services that have been given to the customers. Credit risk has always been an extensively studied area in bank lending decisions. Credit risk plays a crucial role for banks and financial institutions, especially for commercial banks and it is always difficult to interpret and manage. Due to the advancements in technology, banks have managed to reduce the costs, in order to develop robust and sophisticated systems and models to predict and manage credit risk.

Credit Risk Prediction Using Artificial Neural Network
Algorithm

To predict the credit default, several methods have been created and proposed. The use of method depends on the complexity of banks and financial institutions, size and type of the loan. The commonly used method has been discrimination analysis. This method uses a score function that helps in decision making whereas some researchers have stated doubts on the validity of discriminates analysis because of its restrictive assumptions; normality and independence among variables [4]. Artificial neural network models have created to overcome the shortcomings of other inefficient credit default models.

The objective of this paper is to study the ability of neural network algorithms to tackle the problem of predicting credit default, that network algorithm is applied to a small dataset of residential mortgages applications of a bank to predict the credit default. The output of the model will generate a binary value that can be used as a classifier that will help banks to identify whether the borrower will default or not default. This paper will follow an empirical approach which will discuss two neural network-based models and experimental results will be reported by training and validating the models on residential mortgage loan applications. As the final step in the direction, linear regression method is also performed on the dataset.

2 Methodology

2.1 Data

Data has been collected from kaggle.com (lending club loan data) that consists of more than 8.5 million records. A random sample data of 60,000 records have been pulled out from the dataset and appropriate attribute selection has been done from 80 attributes. Attribute selection includes numeric and integer attributes along with some factor attribute relevant to the problem this paper is dealing with. Dataset consists of combination of variables as follows:

1. **Dependent Variable:** loan_status(0 and 1);if the borrower will default then the investment will be bad and if the borrower will not default then he or she will be able to repay the full loan amount. So, to differentiate in neural network 0 indicates borrower will default and 1 indicates borrower will not default.

Welcome to the newly launched Education Spotlight page!

[VIEW LISTINGS](#)

...n_amnt, funded_amnt, emp_length,
...pe

In this study, a classic feed-forward neural network has been used. The feedforward network consists of an input layer with 10 input variables, 7 hidden layers and an output layer with one neuron that represents a classifier. The network is trained by using a supervised learning algorithm(back propagation algorithm. The algorithm optimises the neuron weights be minimising the error between actual and desired output. Error is for neuron i . Weights will be updated by formula, where f be the learning coefficient and is the output from hidden layer. Algorithm will work until a stopping criterion is found.

It is necessary to carefully choose the parameters, such as the value of f and a number of neurons and number of hidden layers, for the neural network algorithm as shown in Fig. 3. In fig 3 connections are represented by black lines between every layer and weights and the blue line shows the bias(intercept of the model) in every step. The network is a black box and training algorithm is ready to use as it is converged. Also, a random sample has been created from the extracted dataset for the network algorithm. Then a training and test dataset is created used to train the model and to validate the performance of the model respectively.

Credit Risk Prediction Using Artificial Neural Network Algorithm

Figure 3: Neural network plot of the credit default model

3 Experiments and Results

There are 10 normalised variables have been fed as the input to the network arranged in an order. The output of the network is a classifier that results in 0 and 1. At first, data has been checked for missing datapoint value, no data was missing; there was no need to fix the dataset. Correlation matrix of the inputs have been shown in Fig. 4.

Credit Risk Prediction Using Artificial Neural Network Algorithm

Figure 4: Correlation Plot of the input dataset

Once the dataset was trained, it was tested on the test dataset. To compute the output based on the other inputs, **compute** function has been used. 7 hidden layers were added to the network and model was created. Following result matrix has been generated by the network:

Table 1: Result matrix for classic feed forward neural network



Error	322.833
Reached Threshold	0.0998
Steps	6765

Total 6765 steps were needed until all derivatives of the error function was smaller than default threshold(0.01). After implementing classic feed forward algorithm, another model has been implemented by using back propagation algorithm with 0.01 learning rate. Classic process and back propagation process have almost same error rate. Thus, classic model fit is less satisfied than back propagation algorithm.

Credit Risk Prediction Using Artificial Neural Network Algorithm

Welcome to the newly launched Education Spotlight page!

[VIEW LISTINGS](#)

ithm

Table 2: Comparison of predicted output and desired output

Actual	Prediction	Matches
0	0.0032	True
0	0.00017	True
0	0.0114	True
1	0.985	True
0	0.0060	True
0	0.0132	True
0	0.9704	False
0	0.0101	True
1	0.00128	True

Last, linear regression have been applied to the dataset in order to compare the accuracy of both the algorithms. glm() function has been used to fit the linear regression model. For regression a probability greater than 0.5 has been assigned, if predicted values in the regression are greater than 0.5, then the value is 1 else 0. Accuracy has been calculated by incorporating misclassification error and confusion matrix has also calculated as shown in Fig. 6.



 Credit Risk Prediction Using Artificial Neural Network Algorithm

Figure 6: Confusion matrix and statistics of linear regression

To highlight the comparison, mean square error of both linear regression and the neural network has been calculated as shown in table 3. As can be seen in the table mean square error of both the process is approximately same and thus both the process are doing same work. It is necessary to know that deviation in MSE depends on the training and test split.

Table 3: Mean Square error of both the processes

MSE neural network	MSE linear regression
0.0220449	0.0227334

4 Conclusion

Welcome to the newly launched Education Spotlight page!

[VIEW LISTINGS](#)

efault. Both the system has been trained
qual effect on the data set and thus are
system classifies the output variable

correctly with a very low error. So, both process can be used to identify credit default with equal accuracy. Also, the neural network represents a black box method such that it is difficult to explain the outcome compared to the linear regression model. Therefore, which model to use depends on the application one has to use. Moreover, while fitting a model using neural network process user needs to take extra care of the attributes and data normalization to improve the performance. To conclude, neural network provides strong evidence to efficiently predict the credit default for a loan application.

Neural network algorithms have a wide range of applications that are not only essential for residential mortgages. Other applications can be rating bonds issued by companies commonly known as bond rating, rating short term investments that can last up to 1 year, long-term and short-term ratings of local and foreign currencies, sovereign, or country ratings. The prediction system can be further enhanced to assign a credit rating to an application by using appropriate algorithms and technologies.

Keywords: Credit Risk, Artificial Neural Network, Linear Regression

About Author:

Shruti Goyal has degrees in MSc Business Analytics from UCD Michael Smurfit Graduate Business School, Dublin, Ireland and B.Tech. from Guru Gobind Singh Indraprastha University Delhi, India. She has over 4 years of experience in data analytics and performance management consulting, previous she has worked IBM Labs and Tata Consultancy Services. Currently she is working as Co-Founder at Reckon Analytics.



Uncategorized

◀ PREVIOUS

[DSC Webinar Series: The Promise of Natural Language Processing \(NLP\)](#)

NEXT ▶

[A Simple Introduction to Complex Stochastic Processes – Part 2](#)

Join us in
ORLANDO
November 5–10

tdwi

New ML, AI, and
Hands-On Python Courses
In-depth training, expert instructors
ML Ops: Best Practices for Delivering

Save 30% w/code **DSC30**

Social media icons: Facebook, Twitter, Pinterest, LinkedIn, Tumblr, GitHub, Email

Related Content



DSC Weekly 17 October 2023

[Scott Thompson](#) | October 17, 2023 at 3:28 pm



Uncharted digital landscapes and the quest for timeless identity

[Michael Peres](#) | October 17, 2023 at 3:09 pm



Internet Of Things (IoT): Application In Hazardous Locations

[Kartik Gandhi, Dr. Suraj Pardeshi](#) | October 17, 2023 at 3:03 pm



The digital evolution in aviation: how big data and analytics are transforming the industry

[Jeremy Bowen](#) | October 17, 2023 at 2:50 pm

Welcome to the newly launched Education Spotlight page!

[VIEW LISTINGS](#)

November 5–10

New ML, AI, and Hands-On Python Courses
In-depth training, expert instructors

MLOps: Best Practices for Delivering and Managing Machine Learning

Save 30% w/code **DSC30**

[About Us](#) [Contact Us](#) [Partner with Us](#) [Advertise with Us](#) [Write for Us](#) [RSS](#)

[Legal](#) [Terms of Service](#) [Privacy Policy](#) [Do Not Sell or Share My Personal Information](#)

© 2023 TechTarget, Inc.

