



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 11, November 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Comparing the Methodology of Bayesian Neural Network with Artificial Neural Network Using Machine Learning Algorithms

¹Mrs.S.Amirtha Rani Jagulin, ²Dr. A.Venmani

¹Research Scholar, Department of Mathematics and Statistics, College of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, India

²Assistant Professor, Department of Mathematics and Statistics, College of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, India

ABSTRACT: Artificial intelligence is worldwide popular now a days. Symbolic artificial intelligence, evolutionary algorithms and machine learning are the branches of AI. BNN and ANN are the branches of machine learning. A powerful data analysis technique BNN and ANN helps us to predict the parameters uncertainty. Artificial neural networks are used in various industries and research fields. ANN is interconnected with neurons with external or internal information that flows in the network. This study will compare BNN and ANN techniques presented in machine learning. The prediction and accuracy of both the networks is discussed on the basis of machine learning algorithms.

KEYWORDS: Bayesian Neural Network, Artificial Neural Network, Artificial Intelligence, prediction, algorithms, machine learning.

1 INTRODUCTION

Artificial intelligence is an interested topic for the last two decades. AI is applied in all the fields such as customer segmentation, fraud detection, computer vision, speech recognition and medical diagnosis [1]. AI helps a machine to acts or thinks like a human by using the censored data. The brain always has its intelligence on its own, AI helps a machine to think, to plan, to solve problems, to act like humans. Symbolic artificial intelligence, evolutionary algorithms and machine learning are the branches of AI. BNN and ANN are the branches of AI. A powerful data analysis technique BNN and ANN helps us to predict the parameters uncertainty. Bayesian inference is based on the prior knowledge, likelihood function and posterior distribution. Bayesian neural network is constructed using the parameters of the posterior distribution. Bayesian networks provide information in cross-sectional data and predicts about the possible effects and interactions among the nodes. Bayesian network behaves like a vehicle for artificial intelligence. Artificial neural networks are used in various industries and research fields. ANN is interconnected with neurons with external or internal information that flows in the network. Machine learning is a subfield of Artificial intelligence that helps the computer to read algorithms and tools from statistics, performs the task without instructions [2].

II. BAYESIAN NEURAL NETWORK (BNN)

Bayesian inference is based on Baye's theorem. Baye's theorem is dependent on the conditional probabilities. The occurrence of one event will depend on the occurrence of the other event is known as conditional probability. The baye's theorem formula is

$$P\left(\frac{A}{B}\right) = \frac{P(A)P\left(\frac{B}{A}\right)}{P(B)}$$

Bayes theorem can interpret the data with uncertainty [3]. In Bayesian, P (A) is a prior event, P (A/B) is the posterior event of A after considering B. P (B/A) is the likelihood function of B for A and P (B) is the sum of all the likelihood function of B for all A [3].



$$P(B) = \sum P(A)P(B \setminus A)$$

Bayesian network is a network model formed using conditional probabilities. This will give a clear picture about the importance or relationship between random variables which are connected using nodes connected through edges. Bayesian inference is based on the prior knowledge, likelihood function and posterior distribution. Bayesian neural network is constructed using the parameters of the posterior distribution. Bayesian networks provide information in cross-sectional data and predicts about the possible effects and interactions among the nodes.

A Bayesian network is a graphical model where each of the nodes represents random variables. Each node is connected to other nodes by directed arcs. Each arc represents a conditional probability distribution of the parents given the children. The directed edges represent the influence of a parent on its children. The nodes usually represent some real-world objects and the arcs represent some physical or logical relationship between them. Bayesian networks are used in many applications like automatic speech recognition, document/image classification, medical diagnosis, and robotics. The BNN will consider the probabilities to form the network and used in the time of uncertainty [4]. In machine learning the term Bayesian network is termed as Naïve Bayes.

2.1 Naïve Bayes

Naive Bayes (NB) is a supervised predictive model known for their simplicity and effectiveness [11]. Naïve Bayesian model is developed from bayes theorem and will help to learn the possibilities of the dataset. The Bayesian inference is based on the prior knowledge, likelihood function and posterior distribution, these models based on the conditional probabilities. The conditional probabilities are independent of the occurrence of features. Given an instance X and a class label θ , using Baye’s theorem, we can express the conditional probability $P(\theta | X)$ as a product of simpler probabilities using the naive independence assumption:

$$P(\theta | f_1, f_2, \dots, f_n) = \frac{P(\theta) P(f_1, f_2, \dots, f_n | \theta)}{P(f_1, \dots, f_n)}$$

Where (f_1, f_2, \dots, f_n) are the features of the vector X .

2.2 Naïve independence assumption

The assumptions on independence probabilities according to naïve bayes we have

$$P(f_i | \theta, f_1, f_2, \dots, f_{i-1}, f_{i+1}, \dots, f_n) = P(f_i | \theta)$$

For all i , we get,

$$P(\theta | f_1, f_2, \dots, f_n) = \frac{P(\theta) \prod P(f_1, \dots, f_n)}{P(f_1, \dots, f_n)}$$

The likelihood function is given as

$$L(\theta) = L(\theta | f_1, \dots, f_n) = P(f_1, f_2, \dots, f_n | \theta)$$

The posterior probabilities are calculated using the function

$$P(\theta | f_1, \dots, f_n) \propto P(\theta) \prod P(f_i | \theta)$$

$$\theta = \operatorname{argmax} P(\theta) \prod P(f_i | \theta)$$

To estimate the probabilities $P(\theta)$ and $P(f_i | \theta)$, Gaussian or normal distribution is used. It is given by

$$P(f_i | \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(f_i - \mu)^2}{2\sigma^2}\right)$$

In real world examples, the naïve Bayesian can be written in the form of

$$P(\theta | f_1, \dots, f_n) = P(\theta | f_1) * P(\theta | f_1, f_2) * P(\theta | f_1, f_2, f_3) \dots * P(\theta | f_1, \dots, f_n)$$

III. ARTIFICIAL NEURAL NETWORKS (ANN)

ANNs are a more efficient tool for predicting and it is useful to examine the reliability and efficiency [5]. ANN is a model of the human brain, which is interconnected with large number of nodes (neurons) working as a system to solve



pattern recognition or data classification problems, particularly in nervous system [5]. ANN is a complex system has an ability to change its structure by adjusting weights of inputs. The neural system is designed to solve problems such as identifying numbered pictures, pictures of cats and dogs etc., this network applies to pattern recognition, optical character recognition, object detection. There neural networks can be classified as

Supervised learning: A pattern which has already knew the output is feed inside the system; the system will recognize the input and give clear output.

Unsupervised learning: An unknown pattern is feed inside the system to get an answer. It will split the data into groups and to identify with the existing data sets.

Reinforcement learning: The ANN takes a decision by recognizing the environment.

Let X_1, X_2, \dots, X_n are the input variables, W_1, W_2, \dots, W_n are weights respective inputs to form the net inputs, b is the bias, which is summed with the weighted inputs to form the net inputs. The parameter of the network are bias and weights, which are adjustable. The output of the neuron ranges between $-\infty$ to $+\infty$. The boundary does not exist so the mapping of input and output should be done. This is known as Activation function.

$$Y = \Sigma (\text{weight} * \text{input}) + \text{bias}$$

IV. DATA DESCRIPTION

The 'Social_Network_Ads.csv' is a data taken from online repository, deals with the purchasing of a particular product through the advertisement. The data consists of 5 columns, they are user ID, gender, age, estimated salary, purchased. Here purchased is a dependent variable and the gender, age, estimated salaries are the independent variable. The purchased column deals with values 0 or 1, 0 represents not purchased and 1 represents purchased.

The proposed model is

$$\text{Purchased} = \alpha_0 + \alpha_1 \text{Gender} + \alpha_2 \text{Age} + \alpha_3 \text{Estimated salary} + \epsilon$$

Considered for experimenting naïve bayes and artificial neural network using the code and predicting the response variable Y (purchased) corresponding to the X variables gender, age, estimated salary. This dataset helps to study about the dependence of purchase of a product on gender, age, estimated salary. The whole data set is divided into training and testing dataset as 75% and 25% for the best prediction. The variable purchased is taken as a factor not a number, assigning 0 or 1 for calculation. The confusion matrix will help us to get a clear picture on prediction.

V. MACHINE LEARNING ALGORITHMS

5.1 NaiveBayes algorithm :

The Naive Bayes algorithm is applied for the data and the result is predicted using the confusion matrix and naïve Bayes graph. The first step is to Import the dataset using the code.

```
dataset = read.csv('Social_Network_Ads.csv')
```

```
dataset = dataset[3:5]
```

The target factor is

```
dataset$Purchased = factor(dataset$Purchased, levels = c(0, 1))
```

Splitting the dataset into the Training set and Test set usig the codes

```
split = sample.split(dataset$Purchased, SplitRatio = 0.75)
```

```
training_set = subset(dataset, split == TRUE)
```

```
test_set = subset(dataset, split == FALSE)
```

The package 'e1071' is an important package for the calculation for naïve Bayes.

```
install.packages('e1071')
```

```
library(e1071)
```

```
nb = naiveBayes(x = training_set[-3],  
y = training_set$Purchased)
```

Predicting the Test set results using and the confusion matrix will be

```
y_pred = predict(naiveBayes, newdata = test_set[-3])
```

```
cm = table(test_set[, 3], y_pred)
```




```
plot(set[, -3],
      main = 'naiveBayes (Training set)',
      xlab = 'Age', ylab = 'Estimated Salary',
      xlim = range(X1), ylim = range(X2))
plot(set[, -3], main = 'naiveBayes (Test set)',
      xlab = 'Age', ylab = 'Estimated Salary',
      xlim = range(X1), ylim = range(X2))
```

5.2 Artificial Neural Network Algorithm

The package 'h2o' is installed to compute artificial neural network

```
# install.packages('h2o')
library(h2o)
h2o.init(nthreads = -1)
model = h2o.deeplearning(y = 'purchased',
                          training_frame = as.h2o(training_set),
                          activation = 'Rectifier',
                          hidden = c(5,5),
                          epochs = 100,
                          train_samples_per_iteration = -3)

# Predicting the Test set results
y_pred = h2o.predict(model, newdata = as.h2o(test_set[-3]))
y_pred = (y_pred > 0.5)
y_pred = as.vector(y_pred)

# Making the Confusion Matrix
cm = table(test_set[, 3], y_pred)

# h2o.shutdown()
```

VI. RESULT AND CONCLUSIONS

The prediction using naïveBayes and artificial neural network is done for the dataset and found that 57 of them does not purchased the product and 29 of them has purchased through the advertisement and 7 has wrongly predicted as purchased out of 100 testing dataset. The total dataset 400 samples are divided into 300 samples for training dataset and 100 for testing data set.

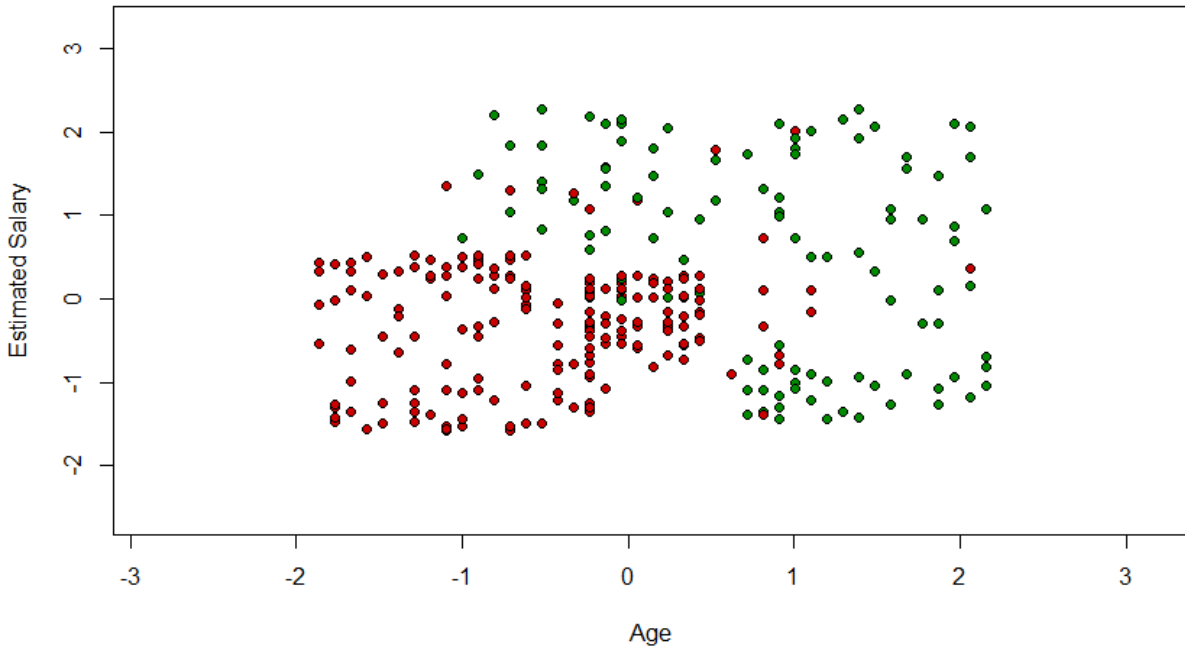
The confusion matrix we got as an output using both the network is same. If we increase the data set it may vary according to the sample size.

	Y pred	
	0	1
0	57	7
1	7	29

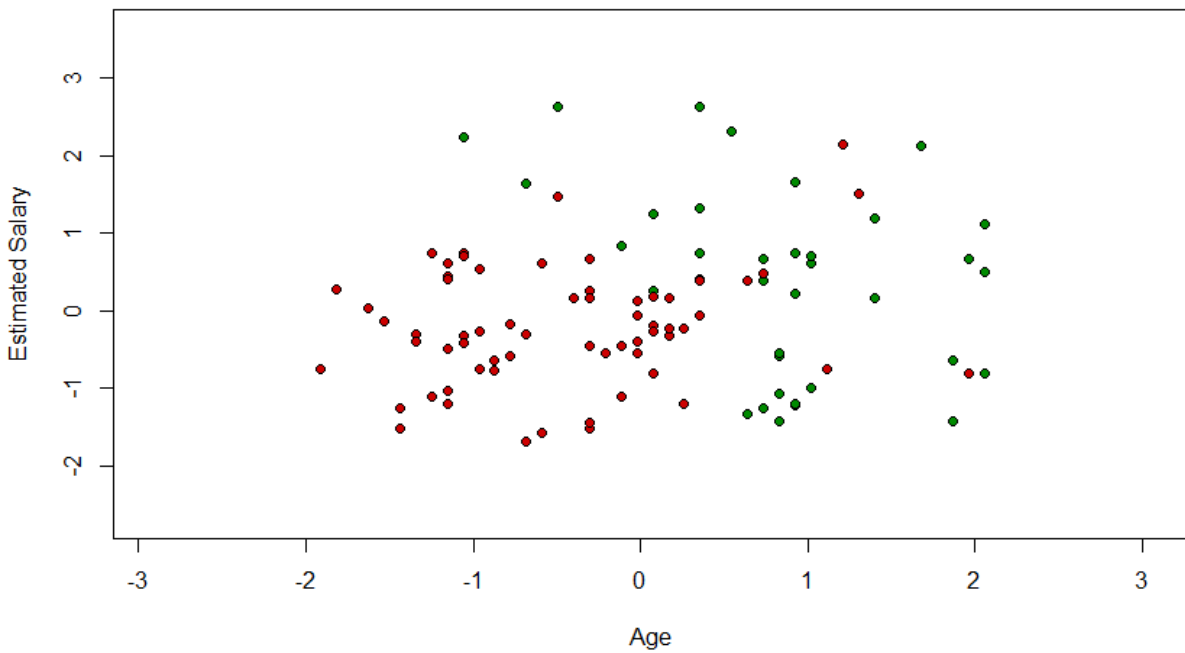
The graphs for training and test data are as follows:



naiveBayes (Training set)



naiveBayes(Test set)



The graphs clearly explain how the points are scattered about the purchase made with respect to age and estimated salary. The graph is made separately for training dataset of 300 samples and test datasets of 100 samples for more clarification.



REFERENCES

1. Shen Liu and Yang Xie, Computational and Statistical Methods for Analysing Big Data with Applications, (2016).
2. Giovanni Briganti and Marco Scutari, An Introduction to Bayesian Artificial Intelligence in Medicine, Université Libre de Bruxelles, IDSIA Dalle Molle Institute for Artificial Intelligence August 7, (2020).
3. Sunghae Jun, Bayesian Inference and Learning for Neural Networks and Deep Learning, Department of Big Data and Statistics Cheongju University Chungbuk, Korea (2020).
4. Ross Harper and Joshua Southern, A Bayesian Deep Learning Framework for End-To-End Prediction of Emotion from Heartbeat IEEE, (2019).
5. Germanno Teles et.al, Artificial neural network and Bayesian network models for credit risk prediction, Instituto de Telecomunicac,oes, Universidade da Beira Interior, Covilh, Portugal,(2020).
6. Marco Scutari, Learning Bayesian Networks with the bnlearn R Package University of Padova, Journal of Statistical Software MMMMM YYYYY, Volume VV, Issue II.(2010) <http://www.jstatsoft.org/>(2009).
7. Patrick mcclure et al, knowing what you know in brain segmentation using Bayesian deep neural networks, frontiers in neuroinformatics, (2018).
8. Allen B. Downey, Think Bayes, Bayesian statistics made simple, green tea press, Massachusetts
9. Tayelor valerio et al., Childless older Americans:2018 current population reports , (2021).
10. Hong chen et al, Improved naïve Bayes classification algorithm for traffic risk management, EURASIP Journal on advances in signal processing(2021).
11. L Ali, et al., A feature- driven decision support system for heart failure prediction based on statistical model& Gaussian naïve bayes, computational & mathematical methods in medicine, (2019).
12. UDEMY, Machine learning A-Z[™] : Hands- on Python &R in Data science (online course).



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.118



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details