

Neural Network Machine Learning Analysis for Noisy Data: R Programming

Yagyanath Rimal, Saikat Gochhait



Abstract: *Neural community machine learning analytical assessment paper discusses the explanatory distinction of secondary data via making use of neural network computing for noisy records the utilization of R programming. Although there's massive hole between the dedication of excellent equipment to analyze overfitting and multicollinearity archives devices for many researchers. Its most necessary aims are to analyzed secondary whose files have been tremendously validation of binary documents base devices of 4 hundred files four variables from internet. The neural community strategies of evaluation are used for prediction whether or not or no longer the university students had been admitted or no with referring their preceding documents the utilization of R software. The outputs with many graphical interoperations had been usually provide a clarification for to gain analytical conclusion. Initially community mannequin with single and more than one hidden layer for first-class admit prediction variable conversed at 11197 for single hidden layer and 5811 for multiple hidden layers. The output and confusion matrixes have been in addition analyzed with developing and reducing of hidden layer that minimized the blunders had been considerably decreased from 33.7 percentage to 23 percentage when the use of rprop+ algorithm and stepmax at one hundred thousand. Therefore, this paper offers best way of computing for inspecting noisy data evaluation when data devices with multicollinearity using r application.*

Keywords : *Hidden Neuron, Over fitted Data, Rectified Linear Unit, Multi-Layer Perceptron*

I. INTRODUCTION

The network analysis is versatile software in the contemporary world, whose design was once exceedingly applicable in automatic automobiles, in the processing of scientific image cognizance and an artificial brain that is based totally on the database algorithm at a time referred to as neural network artificial. The neural evaluation is science of neuronal study as similar to the neural system of the human body, in which each and every neuron is related to dendrites that transmit alerts to some other neuro in the form of electrical impulses (Brains, 2018). Then body passes these indicators and decides with unique actions from central brain. The

artificial neuron is heart of neural network. The activation node every so often takes enter from dendrites and then activation neuron performs a competitive likelihood action (Gurney, 2004). That two multiples enter manner produces and produces a single output where all put alerts are analyzed regularly until the appropriate selection is made. The various phrases of put as variables whose magnitude is viewed as a positive weight denoted as (w) are exceeded to the summary attribute which already has some steady distortion, the bias continually evaluates and produces each real and false judgment. In the wake of procuring zero or an enactment work, the top is determined, which is additionally analyzed with the real forecast with the rate y . This strategy is constantly inspected with parallel; straight out or numerical estimations of the inquiry factors, this device is viewed as single-level observation the utilization of a solitary hub. Basically, there are three elements of the neural network output levels enter levels, and hidden ranges two (Dertat, 2018). Input levels takes enter as two variables ship to output ranges for subsequent process. The middle layer has every one of the forces to investigate records for second layer data sources and delivers yield. Be that as it may, the expressions and its phase as a subscript x (1,2,3,4 ... m) and concealed layer h (1,2,3 ... n) ought to choose all through the model detailing. In the event that a network has several concealed levels, is referenced as profound picking up learning of which improve the perusing cost to the Workstation do what is characteristic for human contemplating by utilizing models taken from past accounts. Therefore, a neural network is a hardware and / or software device modeled after the operation. The activation features threshold bypass characteristic passes 1 if the activation with bios is large than zero, in any other case it passes zero. The sigmoid attribute $(1/1 + e^x)$ is considerably used as activation characteristic for great predicting probability. Another is the hyperbolic digression which depends on the sigmoid of course which goes past - 1 to 1, is like the relapse esteems in which the inclination is more profound and $= (e^x - e^{-x}) / (e^x + e^{-x})$. The amended direct unit (RELU) is the incredible model of the neural system model; it additionally creates 1 if x is + ve; in some other case it produces 0, which is notable because of a considerably less expensive structure and a quicker methodology, anyway requires more prominent two innovation of trials. The favored output workout is completed solely when its errors from the preceding y -values return to the neural function, so that mistakes are minimal in each skip.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Yagyanath Rimal , Faculty of Science and Technology, Pokhara University, Nepal . Email: rimal.yagya@gmail.com

Saikat Gochhait *, Symbiosis Institute of Telecom Management department, Constituent of Symbiosis International (Deemed University), India ,Email: saikat.gochhait@sitm.ac.in

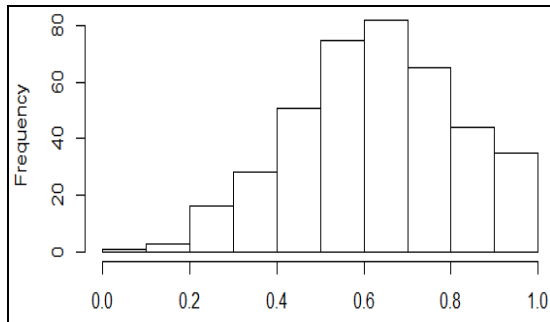
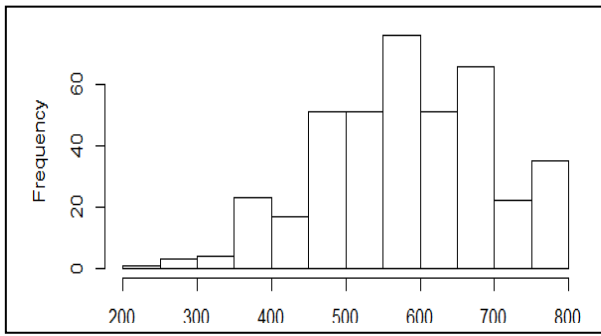
© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The subsequent extension is the enter error comments manner so that its weight is radically reduced for subsequent propagation, so the community should without difficulty have adjusted its weight at each new release so as to reap minimal errors (errors = errors) - 1). $J(\theta) = 1/m \sum_{i=1}^m (z_m - y_i)^2$, this requires extra computing power, use the descent of the systematic gradient which eventually finds the minimum gaining knowledge of rate. The computing device learning process starts with statistics observations, direct experiences or instructions to predict models in the facts and make higher selections in the future. The fundamental intention is to enable computers to research automatically barring human intervention or assistance and to mechanically regulate actions. The giant statistics incorporate the high social, financial and scientific fee that has existed in the course of the educational and industrial world. Big information on actual value, however there are four attributes specific to massive data, regarded as volume, variety, velocity and fee (Dumbill, 2018). Extend the value area of big data. Meanwhile, they develop to be huge challenges for large records analysis. There are three major problems: Big facts from multiple and heterogeneous sources, storage of massive unstructured data and big records of extraordinary value. The three central scientific problems in the contrast of huge information are the representation, archiving and forecast of big data. Traditional techniques can't deal with large data. New techniques for massive information are fundamental (Tsai, 2016). Human brains are, of course, brilliant huge information processors. Neural networks are computational replicas of human data. The history of research in neural networks has ups and downs. Nowadays, with the resource of the enchantment of computational power, the analysis of large facts using neural networks has carried out exceptional success, particularly in huge data applications, for example, large facts analysis, large information analysis, giant scientific assessment data. Neural networks are conducting investigations on synthetic intelligence (Siganos, 2018). The neural networks are capable to extract summaries a couple of sources of information, technique heterogeneous statistics and acquire dynamic adjustments (Zitnik, 2018). They are the scaffold for the usage of gigantic certainties that permit the preparation of neural systems with a monstrous amount of parameters. Nonetheless, there are issues in the neural system model. The size of the network does not have the hypothetical model; and the learning calculation In the colossal records viewpoint, there are three focal logical issues to ensure cognizance in dimensional extra room to speak to the relationship of enormous records estimate (Gudivada, 2017). Neurosciences are vital to solve the integral scientific issues of neural networks and the investigation of giant volumes of data, two evaluation of large amounts of information could be accelerated (Bruke, 2017). The neural system is an equipment and programming program device demonstrated on the working of neurons in the human cerebrum. An engineered neural system models neurotransmitters and natural neurons can furthermore be utilized to make expectations for confused neural systems likewise called manufactured mind systems (Rouse, 2018). The quantity of neurons in the yield layer must be straightforwardly related to the sort of work completed by the neural system. To decide the quantity of neurons in the yield layer, first consider the alleged utilization of the neural network (Saxena, 2018). The quantity of neurons in the information layer relies upon the preparation, given the

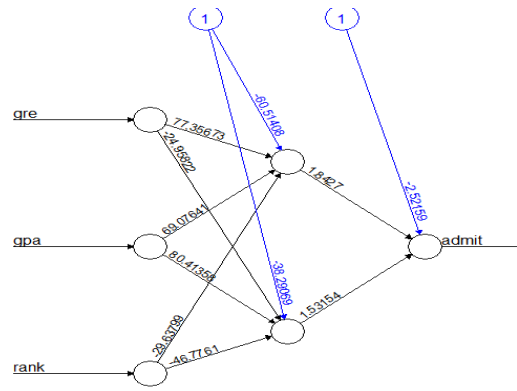
assortment of shrouded layers and the quantity of neurons in each layer, which are typically chosen by methods for a cross-approval system (Bouziane, 2018). According to (Alam, 2015) the most frequent neural network mannequin is the multilayer perceptron (MLP), recognized as a supervised network because it requires preferred output for learning. Neural networks allow a fast and efficient way to model phenomena at different levels of complexity. They can be used in both well-structured and poor problematic processes, whose information is limited. The advantages of neural networks have been transferred to the chances of defining chance distributions of individual activities in cyclical building processes. The additional hidden levels through which errors need to propagate backwards make the gradient greater unstable and the number of false minimums increases. what's more, extreme changes can likewise happen when the instruction units are little in connection to the quantity of shrouded neurons so its preparing set and measurements of concealed layer are consolidated (Sheela, 2013). Albeit dynamic systems can be prepared utilizing the indistinguishable slope based calculations utilized for static systems, the general execution of calculations in dynamic systems can be altogether different and the angle must be determined on a more noteworthy hazardous way.

Neural Network Using R Programming

```
> getwd ()
[ 1 ] "C:/Users/rimal/OneDrive/Documents"
> data= read.csv ("binray.csv", header = TRUE)
> str (data)
"data.frame": 400 obs. of 4 variables
 $admit: internet zero 1 1 1 0 1 1 0 1
 $gre: int 380 660 800 640 520 760 560 400 540
 $gpa: num 3.61 3.67 four 3.19 2 .93 three 2 .98 3.08
 Range of $: internet 3 3 1 4 four 2 1 2 3 2
Here 'm' using binary information frame having four hundred
information with four variables. The admit variable having
properties zero or 1 whether or not student will have admitted
or no longer admitted whose selection is primarily based on
different three variables gre, gpa and rank of student score the
place these former school have been ranked from 1 to four as
Likert scale. While doing neural network the prediction
variable be in categorical and all others values; ought to lies
in-between 0 to 1 this should achieve with the aid of deducting
minimal and dividing max by minimum values.
> hist(data$gre)
>
data$gre=(data$gre-min(data$gre))/(max(data$gre)-min(data
$gre))
>
data$gpa=(data$gpa-min(data$gpa))/(max(data$gpa)-min(da
ta$gpa))
>
data$rank=(data$rank-min(data$rank))/(max(data$rank)-min
(data$rank))
> hist(data$gre)
```



```
> set.seed(#####)
> nd=sample(2, nrow(data), replace=TRUE, prob=c(.7,.3))
> training=data[ind==1,]
> testing=data[ind==2,]
The statistics partitioning creates training data set separate
281 observation of four variables in a similar way checking
out creates 119 statement of four variables. Which is
necessary deal of desktop getting to know the place we format
model from coaching information units and tested with trying
out data units for reliability of prediction n neural network.
> library(neuralnet)
> set.seed(#####)
> nn <-neuralnet (admit~gre+gpa+rank, data=training,
hidden=1,.fcti="ce", linear.output = FALSE)
Here we are growing neural network mannequin with admit
based variable of different three ndependent variable on
training statistics units where hidden neuron s 1 and error time
period s ce and linear output is false.
> plot(nn)
```



Error: 153.914726 Steps: 5811

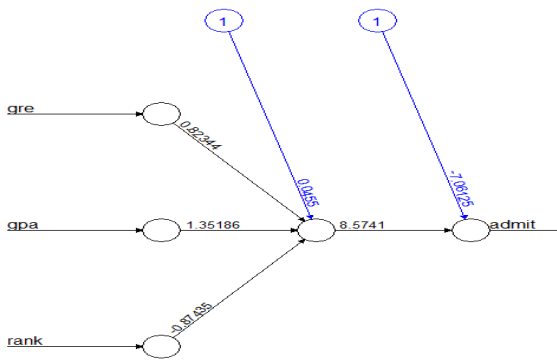
From the above parent three gre, gpa and rank are the predictor variable input layer neuron are absolutely connected with 1 and two hidden layers with ts weight and bias phrases and admit s expected variables whose values had been conversed n 11197 for single hidden layer and 5811 n two neurons n first hidden layers.

```
> output=compute (n, coaching [, -1]) # first column admit
s excluded
> head(output)
neurons" [ [ 1] ]
1 gre two gpa rank
1 two 1 0.275862069 0.7758620689 0.666666667
5 two 1 0.517241379 0.3850574712 1.000000000
6 two 1 0.931034482 0.4252873563 0.333333333
.....
385 1 0.448275862 0.2068965517 0.333333333
387 1 0.896551724 0.9195402299 0.333333333
389 1 0.724137931 0.5229885057 0.333333333
>head (output net. result) [,1]
2 0.377509165 3 0.701123318 4 0.313023265
7 0.336508552 8 0.074358419 9 0.336222113
> head (training [,1])
admit two gre two gpa two rank
21 two 0.758620689 0.810344827 0.666666667
```

Here we get so many outputs the neural network output predict that the 2d teams of records units have 0.337 percent of chance of being admission however normal statistics head of coaching facts categorized as completely admitted s misclassification of model. From the above neural network figure, we can calculate output of every neuron. The hidden neuron input s calculated with sum of bios time period and the product of each enter variable weight and ts input values of three input variables.

```
> n4 <- 0.0455 + (0.82344*0.7586206897) +
(1.35186*0.8103448276) + (-0.87435*0.666666667)
> n4 [1] 1.182751379
```

This s the input to hidden layer whose output is calculated with sigmoid function which is calculated with 1/(1+exponent of preceding input). This method is repeated till final neuron fee is calculated. The returned prorogation again calculated n reversed manner with adjusting ts weight from each closing convergence output.



Error: 161.159929 Steps: 11977



Neural Network Machine Learning Analysis for Noisy Data: R Programming

```
> out4 <- 1/(1+exp(-in4))
> out4 [1] 0.7654421484
> n5 <- -7.06125 + (8.5741*out4)
> n5 [1] -0.4982724753
> out5 <- 1/(1+exp(-in5))
> out5 [1] 0.377946739
```

This price is similar to output of Net. Neuron as above. Confusion matrix and leave out classification errors.

```
> output=compute (n, education [, -1])
> p1=output$net.result
> pred1=if else(p1>=0.5,1,0)
> tab1=table (pred1, training admit)
> tab1
      two      zero      one
pred1 two      zero      one
two    181     57
zero   181     57
one    181     57
```

From the above output values s calculated with the exception of admit variable and end result is saved in some other p1 variable. The prediction is calculated on the basic off the compute output is greater than 0.5 then his/her end result s admitted (1) in any other case non admitted. From the output of confusion matrix 181 individuals were now not admitted expressed via each mannequin and records interoperation the place 57 and 8 information teams have been pass over classified.

```
> 1-sum(diag(tab1))/sum(tab1)
[1] 0.231316726 s omit classification errors
```

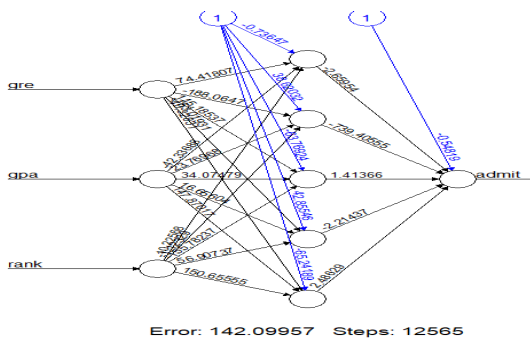
```
Similarly, for test data
> output=compute (n, checking out [, -1])
> p2=output$net.result
```

```
> pred2=if else(p2>= 0.5, 1, 0)
> tab2=table (pred2, testing$admit)
> tab2
      two      zero
pred2 two      zero
two    78     29
zero   78     29
```

Similarly, there's 78+6 s accurate classification however 29+6 had been the omit classification of checking out data

```
> 1-sum(diag(tab2))/sum(tab2)
[1] 0.2941176471 s omit classification errors.
```

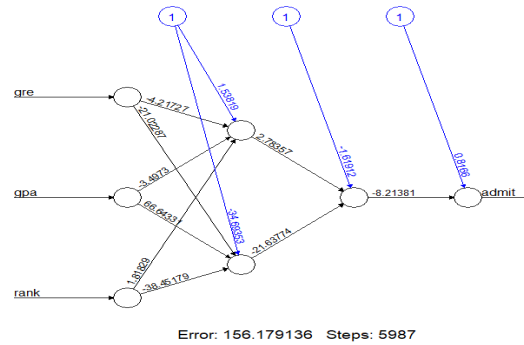
```
Let's put five hidden neuron n the model
> n<-neuralnet (admit~gre+gpa+rank, data=training,
hidden=5, err.fct="ce", linear.output = FALSE)
```



```
> plot(n)
This s completely connected network where each layer nodes were connected with other layers but not connected within each node whose output weight were shown. This network converge n 12565 repetitions whose errors s 142.09.
```

```
pred1 0 1
0 172 53
1 17 39
[1] 0.2491103203
pred2 0 1
0 77 25
1 7 10
[1] 0.26890756
```

The miss classification and errors has decreased than previous model Two hidden neurons



```
pred1 0 1
0 176 62
1 13 30
[1] 0.2669039146
pred2 0 1
0 79 28
1 5 7
[1] 0.277310924
```

The confusion matrix and errors of misclassification shows more worse than above models.

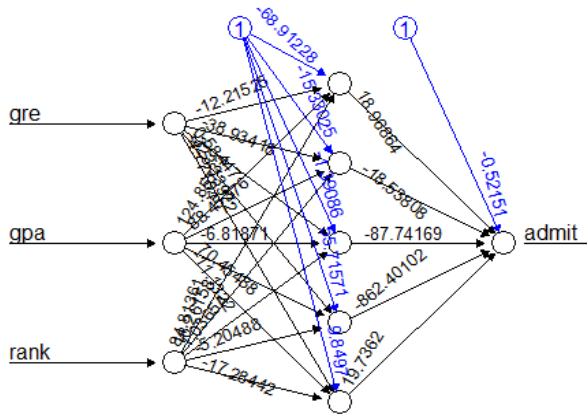
```
With five neurons
>n <-neuralnet (admit ~gre +gpa +rank, data= training,
hidden=5, err.fct="ce", linear.output=false, life
sign="full",rep=5,algorithm="rprop", stepmax=100000)
Output
```

```
hidden: 5 thresh: 0.01 rep: 1 / 5 steps:
1000 min thresh: 0.1286597328
2000 min thresh: 0.1286597328
3000 min thresh: 0.0848351962
17000 min thresh: 0.0107039775
18000 min thresh: 0.0107039775
error: 132.5507 time: 4.81 secs
hidden: 5 thresh: 0.01 rep: 2/5 steps:
Could not converge
Hidden: 5 thresh: 0.01 rep: 3/5steps:
1000 min thresh: 0.092244246
2000 min thresh: 0.092244246
13000 min thresh: 0.010879176
error: 139.80883 time: 3.4 secs
hidden: 5 thresh: 0.01 rep: 4/5 steps:
1000 min thresh: 0.1472573813
2000 min thresh: 0.1472573813
3000 min thresh: 0.09138904351
13000 min thresh: 0.01208854108
14000 min thresh: 0.01091560314
error: 147.41304 time: 3.9
secs
```

```
hidden: 5 thresh: 0.01 rep: 5/5 steps:
1000 min thresh: 0.1131700352
2000 min thresh: 0.102054191
9000 min thresh: 0.01150909692
9000 min thresh: 0.01150909692
10000 min thresh: 0.01012525328
error: 145.64415 time: 2.6
```

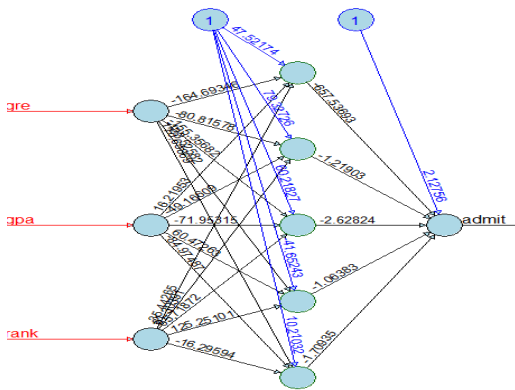
From the above output except 2 nd reputation all network was conversed with step with different error and time taken, this could conclude that the first neuron s best suited because of error: 132.5507 time: 4.81 secs least

```
> plot(n,rep=1)
```



Error: 132.550704 Steps: 18339

```
>plot(n, col.hidden='dark green',col.entry.synapse
='red',show.weights =T,information =F, fill='light blue')
```



```
> output=compute(n,training[,-1],rep=1)
> p1=output$net.result
> pred1=if else(p1>=0.5,1,0)
> tab1=table(pred1,training$admit)
> tab1
pred1 0 1
0 184 67
1 5 25
> 1-sum(diag(tab1))/sum(tab1)
[1] 0.256227758
> #for test data
> output=compute(n,testing[,-1],rep=1)
> p2=output$net.result
> pred2=ifelse(p2>=0.5,1,0)
> tab2=table(pred2,testing$admit)
> tab2
pred2 0 1
0 79 23
```

```
1 5 12
> 1-sum(diag(tab2))/sum(tab2)
[1] 0.2352941176
> n <-
neuralnet(admit~gre+gpa+rank,data=training,hidden=5,err.f
ct="ce",linear.output=FALSE,
lifesign='full',rep=5,algorithm="rprop+",stepmax=100000i)
```

II. CONCLUSION

In this work, two first studied the neural neighborhood algorithm for information analysis. Simulation of neural networks n binary facts units of 4 hundred records. Despite the fact that the neuronal system has a great deal considerably less interpretability than the determination tree. In any case, it is more noteworthy fitting for loud datasets for unrecognized examples. An R regression assessment used to be as soon as performed to measure the correlation between outcomes and objectives. We create a regression graph to validate the network that shows the relationship between network outputs and goals. it is determined that the value of R s nearer to 1. Neural networks are generally used in pattern attention due to their practicable to reply to unexpected locations / patterns. During training, neurons are taught to apprehend one of a kind special patterns and to be or now not shoot when that fashions received. The rectified linear unit (RELU) is the fantastic model of the neural system model; it also delivers 1 if x is +ve; something else, t produces 0, which is popular because of a less very estimated structure and a quicker methodology, anyway requires bigger age of trials. The output and confusion matrix have been analyzed in developing and decreasing of hidden layer that minimized the errors have been considerably diminished from 33.7 percent to 23 percent.

REFERENCES

- Bouziiane, A. (2018). Neural Network Analysis. online.
- Bruke, j. (2017). Big Data Analysis Using Neural Networks. online.
- Gurney, ok. (2004). An introduction to neural network. Ucl press confined s an imprint of the Taylor Francis group. online.
- Haluk, D. (2015). Improvements with smart company systems: Analytics, big data, cognitive help, and the internet of everything. online.
- Johnson, S. (2004). Neural Coding strategies and mechanisms of opposition. cognitive systems studies. online.
- Lim, S. (2018). Adaptive gaining expertise of rule for hardware based totally completely deep neural networks. online.
- McAteer, M. (2018). An introduction to probabilistic pro- gamming, now reachable n TensorFlow probability. online.
- Noon, H. (2013). Artificial neural network: Beginning of the ai revolution. online.
- Rouse, M. (2018). Big information evaluation the use of neural networks. online.
- Saxena, S. (2018). Becoming human: Artificial Genius magazine. online.
- Shwe, M. (2018). An introduction to probabilistic programming, now accessible n TensorFlow probability. retrievedfromhttps://medium.com/TensorFlow/an-introduction-to-probabilistic-programming-now-available-in-TensorFlow-probability. online.
- Zhang, Y. (2017). Big statistics contrast the usage of neural networks. Online.

AUTHORS PROFILE



Yagyanath Rimal with Master Degree in Information Science 2006 with three books published on programming language for students.



fellowship 2019

Saikat Gochhait with Post Doctoral Fellow from University of Extremadura, Spain and with more that 50 publications in reputed journals indexed in scopus and IEEE member and awarded with MOFA Taiwan