DEVELOPING AN EFFICIENT METHOD FOR 'ESTIMATION OF COST' IN CLOUD COMPUTING USING NEURAL NETWORK

Pranshul Pahwa

ABSTRACT

Objectives: To study the cost estimation using FFNN (Feed Forward Neural Network) and BPNN (Back Propagation Neural Network). Methods/Statistical Analysis: In proposed work, resource allocation has been done in which cost function has been estimated. The whole simulation is being done using MATLAB 2010a environment. Findings: From the simulation results, it is analysed that using FFNN, 95% of accuracy is achieved. Application/Improvements: With the advent of this technology, the cost of computation, application hosting, proper storage of content and delivery is abridged considerably.

1. INTRODUCTION

Cloud computing is the new technique for the provision of resources in the network using the concept of virtualization1. It provides the platform for Saas, Iaas like services in less time2. Ever more companies are proposing services of cloud computing as apparent by way of the growth and enlargement of commercial cloud infrastructures, for example, Amazon, ENKI Prima Cloud Gogrid, Flexiant, Layered Technologies, Microsoft, as well as vCloudExpress. According to Google's research, five key features of cloud-computing are consumer centric, intelligence, task centric, powerfulness, and programmability3. Because of the on-demand availability of the resources, cloud computing has become the best model so far. The main focus in this field is given to the domain of cloud computing with allocation as well as re-allocation of the resources. This is for fulfilling those demands that are unpredicted and in return, high investment is achieved4. Therefore, cloud computing has made the applications of business more movable and concerned. Figure 1 shows the usage of private cloud in India during the year 2014- 2015.

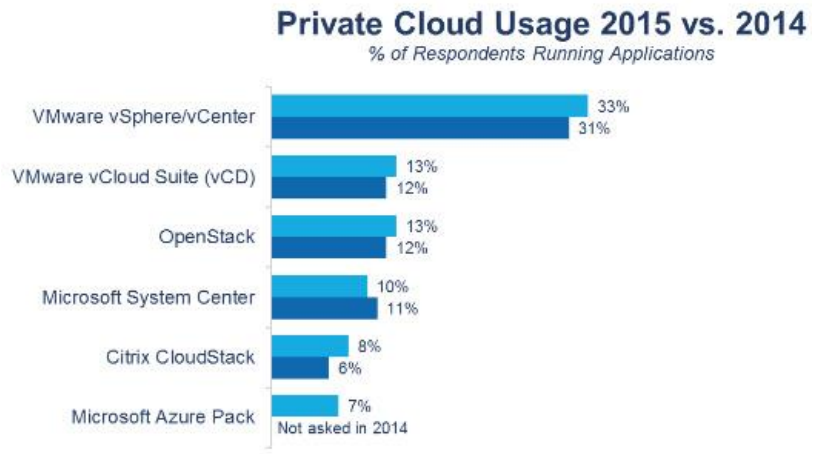

Figure 1. Usage of private clouds.

In this research paper, we proposed a simulation model based on an allocation of resources or task to do the work by reducing the complexity.

1.1 Resource Allocation Significance

Resource allocation is a method which is executed in a number of regions of computing, like data_center managing, operating_systems, as well as grid_computing5. It has the activities scheduling with the resources that are accessible at less cost in the best manner. It executes best resources for developer applications by lessening the cloud environment amount. The energy consumption is linked via allocation of resources that has to be considered6. Resource allocation is the technique is which resources are allocated to the servers that are idle. In this resources are allocated from busy servers to idle servers to reduce the delay, bandwidth etc7.

PCs can be made less vitality utilization gadgets by utilizing lower power processor, utilizing cooling gadgets just as utilizing turning SSD (Solid State Drive) of little size as opposed to huge size8. Intel has built up a procedure that is going to utilize less force. It uses assets like server farms, heat, light, power and so on in an ideal way. Indian IT industry has seen enormous development in the world9. Notwithstanding, this colossal development prompts the high utilization of vitality just as force that outcomes in significant expense utilization. As indicated by measurements from IBM, it shows that under 4 percent of the vitality going into a server farm is utilized to process undertakings. Due to the vast population, natural resources are already on the verge of the extinction10. Due to this, IT companies are moving in the direction of Green Computing as it leads towards decrease of the expenses, and decreases the utilization of the natural resources11. Also Power costs form a major part of the complete operational expense for IT companies. Henceforth, it has become almost essential for IT firms to adopt green computing measures12. The main aspects of resource allocation in accordance with security in the cloud are that specific data which need to be safeguarded while in idle state, in transit state, as well as in use state, and access to the data must be controlled, that is to say as shown in Figure 2:

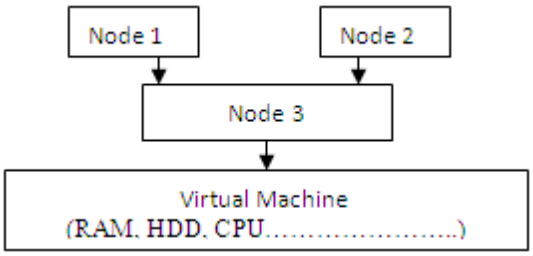

Figure 2. Resource allocation via VM.

To ensure the high level security it is very important to store the cloud data at specific places and handle that data properly13.

Mathematically Resource allocation can be seen as below;

Traffic for single server is shown below

1.2 Resource Allocation Problem

Resource allocation is the type of the problem that arises since the need of resources occurred14. The resources in the cloud are not used directly but they are used via some application like SOAP, UDDI etc. They are the web APIs that are needed for utilization of resources15. As cloud computing has the capability of providing the resources on demand so sometimes resource allocation is not done properly.

The main challenge of cloud computing is the allotment of number of requests of the resources16. These resources are provided so that the QOS can be enhanced and this allocation should be done well.

1.3 Limitations of Resource Allocation in Cloud

• Problem in the resource allocation is when the rights of resources occur. They can only rent the resources.

• For improved storage space of information as of one provider to another is not easy task due to the large amount of data.

• In public cloud, the data of users is prone to various attacks as servers are interlinked with each other, so security issues in resource allocation is common.

• With utilization of cloud, the external devices like printers etc. do not work well. They require

software to be installed at your end.

• More understanding is required for allocation of the resources. As indicated by Amazon's estimations, the energy related expenses for data centre record for 42% of the aggregate working expense. Therefore, it is critical to comprehend the energy proficiency of this rising MapReduce standard and to attempt to lessen the energy utilization in cloud data-centres. There are numerous market oriented methods which leads to the regularity in the power and resource supply. To ensure the robustness of the cloud, they must be adaptable to VM machines by keeping the customers detached from general data centre. VM machines coordinate with each other by allocation the resources based on SLA violation17.

Cloud computing permits hosting of numerous services on a worldwide shared-resource-pool where resources are assigned to several services according to their demand. It utilizes virtualized condition for working administrations, in light of the fact that without virtualization registering is wasteful and not adaptable. Be that as it may, it has some presentation debasements of administrations and furthermore has vitality overheads and huge measure of intensity utilization as appeared in Figure 3.

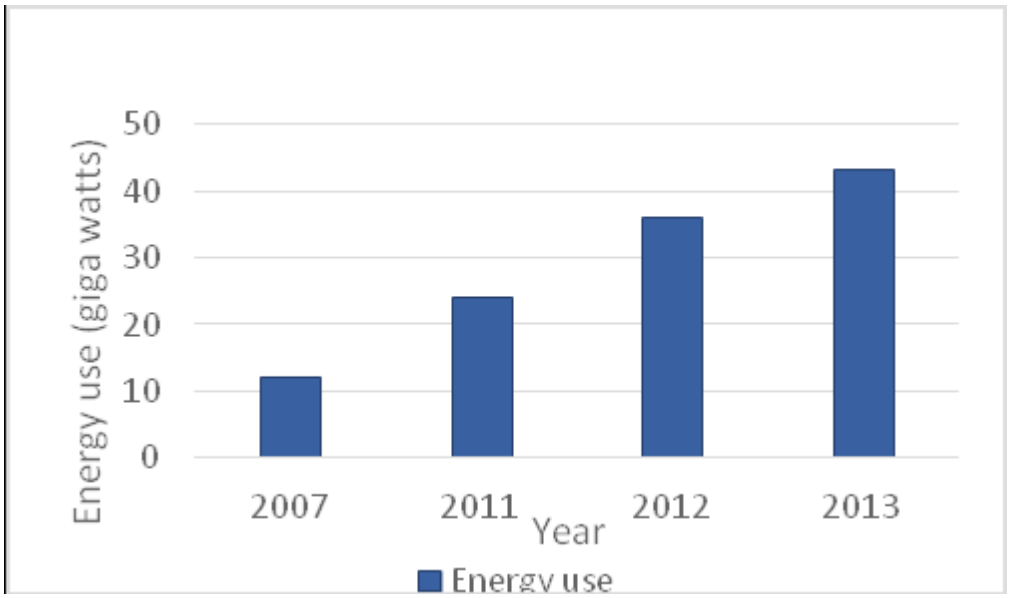

Figure 3. Energy use in global data centres27.

In the past, numerous scientists took a shot at making vitality productive calculation for lessening vitality utilization .Numerous calculations were executed for sparing vitality of server farms by killing or by putting inert servers to rest mode. Be that as it may, these procedures were not all that successful in light of execution corruptions of administrations and inappropriate assets usage. Likewise, it prompts the high utilization of vitality. So it must be dealing with while dispensing the assets as well. A portion of the past works gives a plan to making vitality effective calculation for server farms. The proposed calculation of Virtual machine arrangement for the Minimization of Migration (MM), which thinks about the use of host CPU. The presentation of calculation is superior to other position calculations yet they didn't think about SLA parameters while choosing VM for relocation, which may be animated utilizing live-movement.

The majority of the infringement happens all through the procedure of live movement of virtual machines, relocation sways the parameters of SLA (like accessibility, reaction time, throughput, organize transmission capacity and so on). Along these lines, there is a need to grow new methodology for SLA mindful vitality effective calculation for asset portion in server farms. This paper has introduced the near audit on vitality minimization utilizing VM relocation just as SLA infringement.

1.4 Virtual Machine

The possibility of Virtual Machines (VMs) is associated with lessening the vitality use as it basically diminishes the pace of inactive force in the general base and its working has appeared in Figure 4.

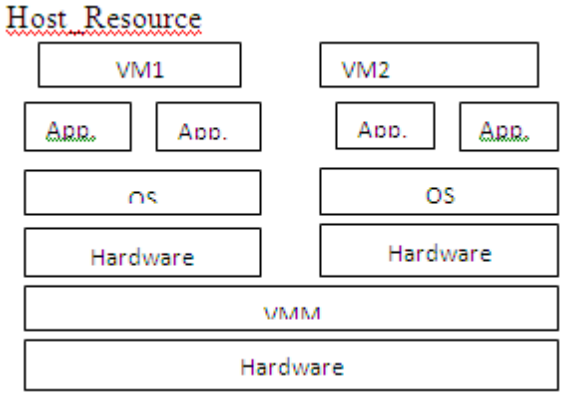

Figure 4. Virtual machine migration48.

Suppose there are large number of resources varies from R1……..Rn in the pool of data centre. Then allocation of these resources to virtual machines (V1,……Vn) will be done as R1V1,……RnVn) and mathematically it can be written as;

$$PS * \sum RC * \frac{\sum VT}{\sum R}$$

Since its beginning, the IT industry has focused on the deployment of the IT equipment and services. Therefore, the main emphasis has been on the processing power and systems. Less attention was given to data centres resulting in high energy consumption and more space requirements for data centres. Investment in data centres including energy cost and cooling them. A large enterprise data centre costs from $506K to $2B (Google data centre). One study indicates that the high cost of running data centre is due to high energy requirements for cooling the equipment's. The high use of the electricity in the data centres makes the use of green computing logical in optimizing resources18.

1.5 SLA Violation

Server consolidation is mostly utilized to maximize the usage of available resources. Presently, the main method of server consolidation is virtual migration. Conventional methodologies, as a rule, decide a special VMs position before sending help in the generation condition. In any case, such methodologies are not appropriate to cloud framework because of the heterogeneity and the time-fluctuation of the outstanding task at hand under which facilitated administrations are oppressed, which cause changes in the usable states of the administrations and if not mulled over, may build the likelihood of an SLA infringement.

2. A GLANCE OF EXISTING TECHNIQUES

Allocation of resources has a major role in the field of cloud computing, frequently in this field, the users distributes data in clouds similar to application etc., other than the virtualization users that share the communications.

2.1 For resource Allocation

In19, author utilised the feedback algorithm for the management of virtualised machines. In this VM machines are grouped together into shared pool then as per SLO agreement, the VM allocation will takes place. In20, two- layer model of utility functions has been proposed for resource allocation that consists of agents and global provider.

In21, resource allocation method based on pre-emptive tasks is proposed for task executions in cloud computing. In22, HPC based resource allocation is presented in cloud computing to investigate the need of resource allocation. In23, SLA utilization for resource allocation is being done.

In24, reduction of energy as well as resource utilization has been done.

2.2 For Energy Efficiency

In25, dynamic resource allocation based on virtual migration concept is being done for ensuring the QOS and to minimise the power consumption using genetic algorithm as well as MBFD method. In26, presented the method for CPU utilization based on threshold based method to balance the workload. In27, round robin that is dynamic in nature is proposed using virtual machine concept for power consumption. In addition to this first- fit algorithm is also proposed. In28, a new algorithm called OBFD is proposed for energy, power with the allocation of resource in cloud computing.

3. PROPOSED WORK MODEL

Neural network utilization has been done for resource utilization in cloud computing as shown in Figure 5. Following algorithms will be used to run the proposed model;

Input parameters:

```
data=xlsread('workdiary.xlsx');
mywork{1,1}='WORK ID';
mywork{1,2}='WORK TIME';
mywork{1,3}='WORK PRICE';
mywork{1,4}='WD-P/T';
workid=data(:,1);
worktime=data(:,2);
workprice=data(:,3);
workd=data(:,4);
[r,c]=size(data);
```

Algorithm 2: Gather Vendor Information

```
data=xlsread('vendors.xlsx');
[r,c]=size(data); % the amount of rows in the xlsfile
would represent the total number of vendors in the list
vendorcount=r;
paid=data(:,1);
param=data(:,2);
paprocessor=data(:,3);
pp=data(:,4);
myvendor{1,1}='Vendor ID';
myvendor{1,2}='Vendor RAM';
myvendor{1,3}='Vendor PROCESSOR';
myvendor{1,4}='RP REL';
fori=1:vendorcount
myvendor{i+1,1}=paid(i);
myvendor{i+1,2}=param(i);
myvendor{i+1,3}=paprocessor(i);
parel(i)=paprocessor(i)*param(i);
myvendor{i+1,4}=parel(i);
vendorinwork(i,1)=paid(i);
vendorinwork(i,2)=param(i);
vendorinwork(i,3)=paprocessor(i);
vendorinwork(i,4)=parel(i);
vendorinwork(i,5)=pp(i);
end
```

Algorithm 3: Work Vendor Information

```
[vendorcount,propcount]=size(vendorinwork);
[workcount,workcol]=size(workdescription);
workpool=[];
wjobappcount=0;
wpdescription=[];
wpdescription{1,1}='VENDOR ID';
wpdescription{1,2}='WORK ID';
wpdescription{1,3}='WORK DENSITY';
wpdescription{1,4}='VENDOR DEMAND';
fori=1:vendorcount
vendormrprelation=vendorinwork(i,4);
for j=1:workcount
currentworkdensity= workdescription(j,4);
currentworkdays=workdescription(j,2);
applicationa=vendormrprelation/currentworkdays;
ifapplicationa<currentworkdensity
wjobappcount=wjobappcount+1;
workpool(wjobappcount,1)=i; % I IS VENDOR

workpool(wjobappcount,2)=j;  % J IS THE JOB
workpool(wjobappcount,3)=currentworkdensity;
workpool(wjobappcount,4)=applicationa;
wpdescription{wjobappcount+1,1}=i;
wpdescription{wjobappcount+1,2}=j;
wpdescription{wjobappcount+1,3}=currentworkden
sity;
wpdescription{wjobappcount+1,4}=applicationa;
end
end
```

Algorithm 4: Training using BPNN network

```
load applications
maximumjob=max(myjobid);
totaltrainingrows=1;
fori=1:maximumjob
totaltrainingrows=1;
Trainingdata=[];
group=[];
 [s,pos]=find(myjobid==i); % places where current
job is placed
for j=1:numel(pos)
Trainingdata(totaltrainingrows)=mycost(pos(j));
group(totaltrainingrows)=j;
totaltrainingrows=totaltrainingrows+1;
end
try
net=newff(Trainingdata,group,10); % architecture of
bpnn ---- newff(Trainingdata,target,hidden neurons)
net.trainparam.epochs=50;
net=train(net,Trainingdata,group);
disp(strcat('JOB NO :',num2str(i)));
catch
disp('no applicant found');
end
end
```

Algorithm 5: Sigma function of neural network

```
function [net test]= trainBPNN(x)
x=double(imresize((x),[8 8]));
net=newff(x,x,[25,5,25],{'tansig','tansig','tansig'},'trai
nlm');
net.trainParam.epochs = 100;
net = train(net,x,x);
test=sim(net,x);
test=round(test);
return;
```

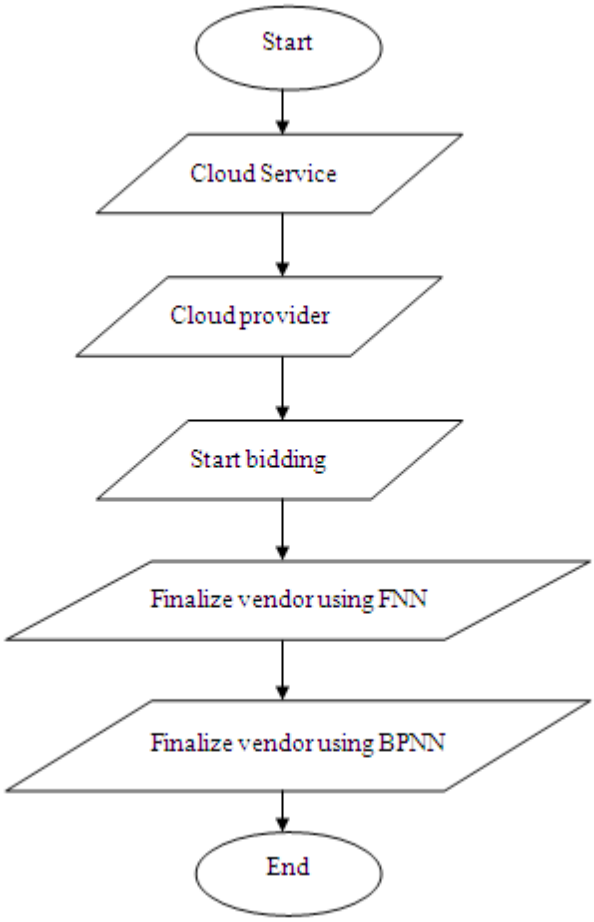

Figure 5. Flowchart of work.

4. RESULTS AND DISCUSSION

To verify the effectiveness (qualities and performance of system) of the proposed method, we conduct several experiments with this procedure on data. There are some steps of our proposed technique are given below:

Phase 1: Firstly, we develop a particular GUI for this implementation. After that we generate work with different ID, time and price for different work.

Phase 2: Develop a code for the cloud server providers which provide a vendor.

Phase 3: After that start bidding and estimate the cost of work according to the data.

Phase 4: Lastly we finalise the vendor according to appropriate work using feed forward neural network as well as by BPNN. After that we calculate parameters like Accuracy and Error Rate and compare them with both approaches.

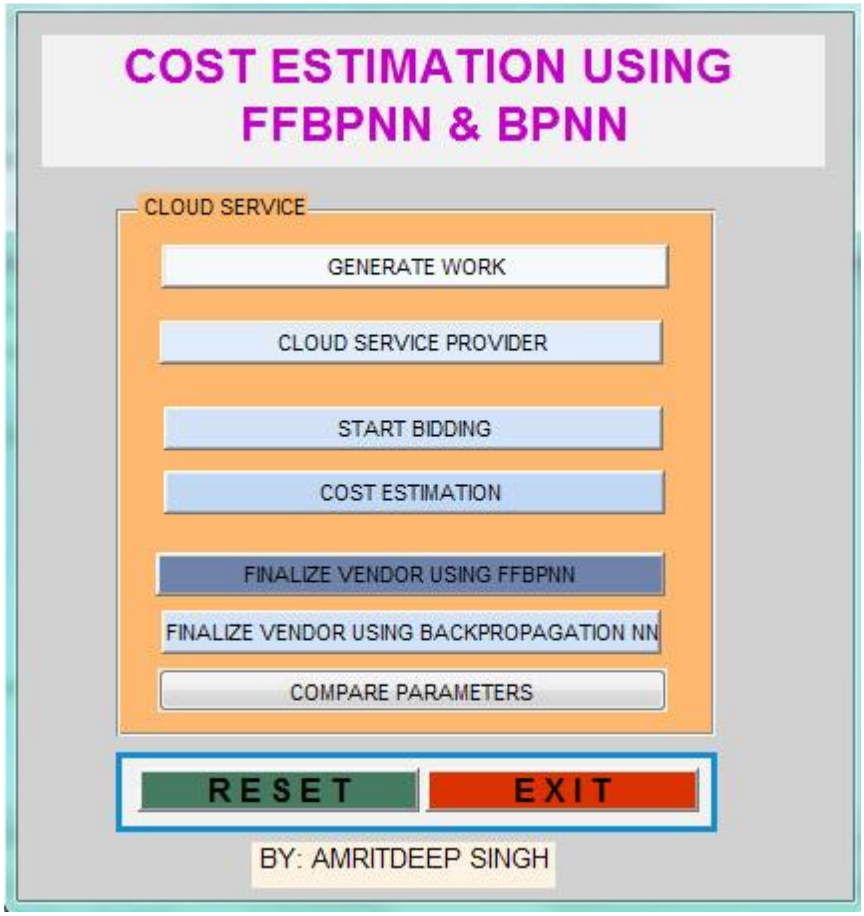

Figure 6. Working window for proposed system.

Figure 6 shows the working window in which we operate our proposed system step by step. The system uses feed forward neural network as well as by BPNN.

Training is being done via FBPNN method as shown in Figure 7. It is like large neuron in which every unit is connected with all units in the layers. The weights of these networks contain the knowledge. Generally, units concerned with network are also called nodes. Data centre takes input and the then passed it over network. During normal operation it worked as classifier but sometimes it acts as like feedback between layers. So, it named called FBNN.

Figure 7. Training through FBPNN.

BPNN is the acronym for back propagation neural network in which to adjust the weights each unit has to propagate in backward direction as shown in Figure 8. Various functions have been utilized in BPNN like no. of epochs, gradient function and validation check number.

It is considered as the supervised learning method.

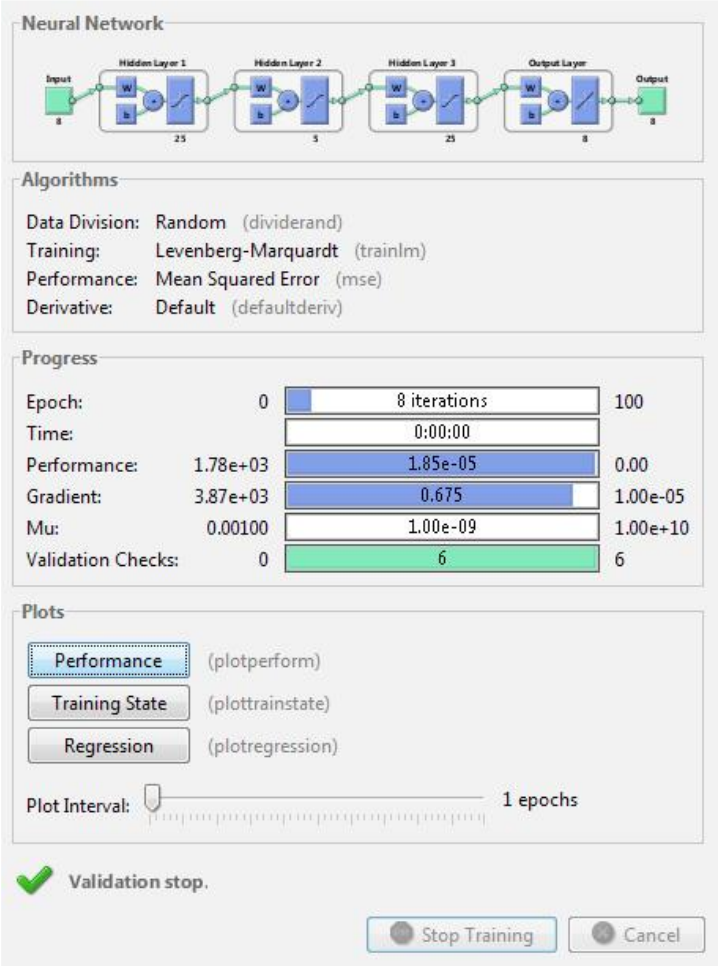

Figure 8. Training through BPNN.

Comparison of Error rate with the proposed system and FFNN and BPNN is shown in Figure 10. It has been shown that the error rate in FFNN is less as compare to BPNN.

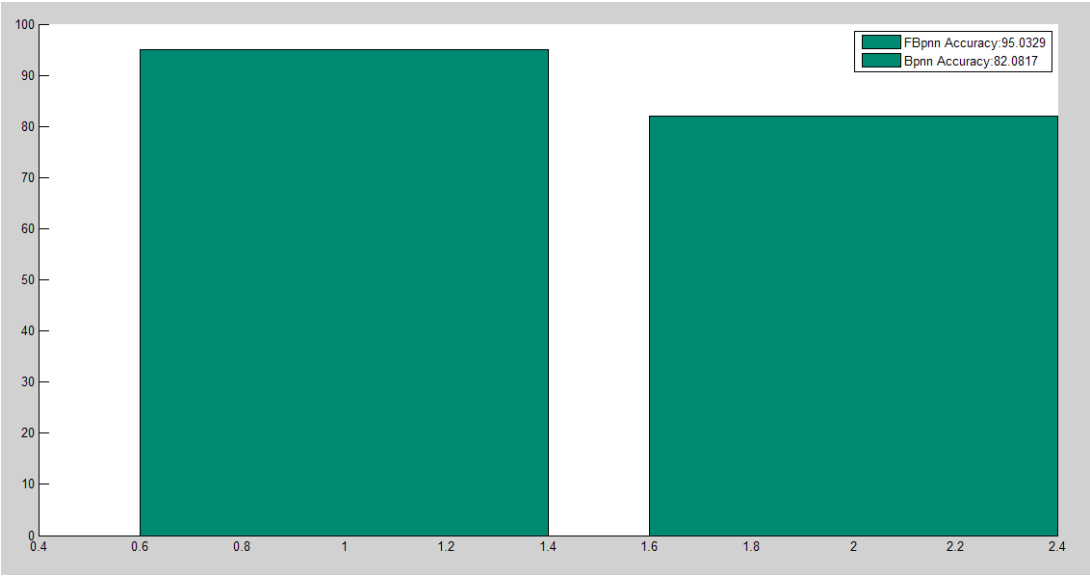

Figure 9. Comparison of accuracy.

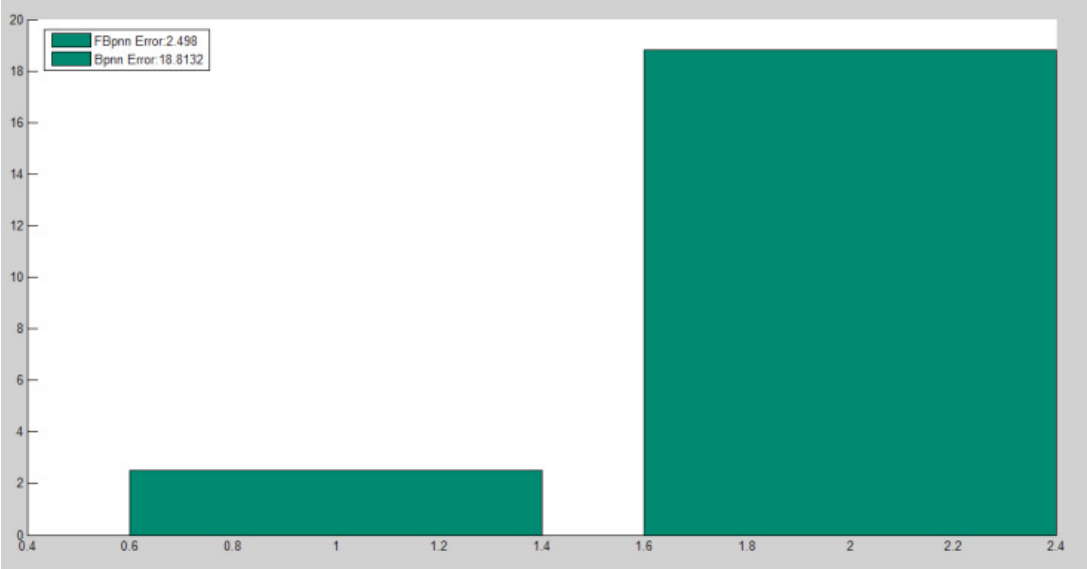

Figure 10. Comparison of error rate.

5. CONCLUSION

For the unpredictable demands, allocating and de-allocating of the resources is done dynamically in cloud computing that in the end, provides better investment. A system is proposed that differentiate the cost among the techniques and to find the system with less error rate. As per the scenario, the cost of FFNN came out to be best as compare to BPNN in cloud computing. In case of FFNN, the accuracy came out to be 95% as compared to the case of BPNN, when the accuracy is 80%.