

A STUDY OF ACCURACY OF NEURAL NETWORKS, SUPPORT
VECTOR MACHINES AND THE BAYESIAN REGRESSION IN
LOAD DEMAND AND PRICE FORECASTING



By

Talha Umair Sultan

Bahria University Islamabad Campus - Pakistan

CERTIFICATE OF ORIGINALITY

I certify that the intellectual contents of the thesis

“A Study of Accuracy of Neural Networks, Support Vector Machines and the Bayesian Regression in Load Demand and Price Forecasting”

is the product of my own research work except, as cited properly and accurately in the acknowledgment and references, the material taken from any source such as research papers, research journals, books, internet, etc solely to support, elaborate, compare and extend the earlier work, Further, this work has not been submitted previously for a degree at this or any other University.

The incorrectness if the above information, if proved at any stage, shall authorize the university to cancel my degree.

Signature: _____ . Dated: _____ .

Name of the Research student: _____.

**“A Study of Accuracy of Neural Networks, Support Vector
Machines and the Bayesian Regression in Load Demand and
Price Forecasting”**

A Thesis Presented to

Bahria University Islamabad

In partial fulfillment of the requirement for the degree of

MS (Electrical Engineering)

By

Talha Umair Sultan
Enrollment Number 01-244132-046

Spring, 2015

Final Approval

This thesis titled

A Study of Accuracy of Neural Networks, Support Vector
Machines and the Bayesian Regression in Load Demand and
Price Forecasting

By

Talha Umair Sultan
01-244132-046

Has been approved

For the Bahria University, Islamabad

External Examiner: _____
Dr. Shahid Khan

Supervisor: _____
Jehanzeb Ahmed
Department of Electrical Engineering, Bahria University Islamabad Campus

Internal Examiner: _____
Dr. Najam-ul-Islam
Department of Electrical Engineering, Bahria University Islamabad Campus

HoD: _____
Dr. Muhammad Ali Shami
Department of Electrical Engineering, Bahria University Islamabad Campus

Declaration

I, Talha Umair Sultan, (01-244132-046) hereby declare that I have produced the work presented in this thesis, during the scheduled period of study. I also declare that I have not taken any material from any source except referred to wherever due that amount of plagiarism is within acceptable range. If a violation of HEC rules on research has occurred in this thesis, I shall be liable to punishable action under the plagiarism rules of the HEC.

Date: _____

Signature of Student:

Talha Umair Sultan
01-244132-046

**Certificate of Completion
of
Thesis Work**

This is to certify that **Talha Umair Sultan, (01-244132-046)**, has successfully completed his research thesis, titled **A Study of Accuracy of Neural Networks, Support Vector Machines and the Bayesian Regression in Load Demand and Price Forecasting**, under my supervision. The thesis meets the scholarly standards as set by Bahria University, Pakistan.

Date: _____

Supervisor

Jehanzeb Ahmed
Associate Professor,
Dept. of Electrical Engg.
BUI, Islamabad

Dedicated to
My Mom, Dad, Sister
and
My Teachers

Acknowledgements

Thanks to

Almighty Allah,

For giving me the strength, motivation and ability to choose the right path.

My Parents,

To the strongest two people I know, if I were even ten percent a person that you guys are, I would be doing the world a favor.

Mr. Jehanzeb Ahmad,

We did it. *Bravo!*

~Talha

Table of Contents

CHAPTER 1.....	14
INTRODUCTION.....	14
1.1 General:.....	14
1.2 Forecasting Techniques:	15
1.3 Report Structure:.....	17
1.4 Summary:	18
CHAPTER 2.....	19
BACKGROUND AND LITERATURE REVIEW.....	19
2.1 Literature Review:.....	19
Economic factors:.....	19
Time factors:.....	20
Weather factors:.....	20
Random Factors:.....	23
Fuel cost issues:.....	23
2.2 Summary:	26
CHAPTER 3.....	27
REGRESSION.....	27
3.1 Introduction:	27
3.2 Regression:	28
3.3 Types of Regression:	29
3.3.1 Linear Regression:	29
3.3.2 Non-linear Regression:.....	30
3.3.3 Logistic Regression:.....	31
3.3.4 Stepwise regression:.....	32
3.3.5 Generalized Linear Model:	33
3.3.6 Generalized linear mixed models (GLMM):.....	33
3.4 INTERPOLATION AND EXTRAPOLATION:.....	33
3.5 Summary:	35

CHAPTER 4.....	36
METHOD AND EXPERIMENTAL SETUP.....	36
4.1 Aim of the project:	36
4.2 Framework Design:	38
4.3 Data Collection:	40
4.4 Classifier Methods:.....	42
4.4.1 Neural Networks (NN):	43
4.4.2 Support Vector Regression (SVR):.....	45
4.4.3 Bayesian Regression:	50
4.5 Summary:	51
CHAPTER 5.....	52
IMPLEMENTATION AND RESULTS.....	52
5.1 Results using NN:	52
5.2 Results using SVR:.....	57
5.3 Results using Bayesian Regression:	60
5.4 Comparison Plot:.....	62
5.5Summary:	62
CHAPTER 6.....	63
CONCLUSION.....	63
6.1 Discussion:	63
6.2 Conclusion:.....	63
6.3 Future Work:.....	64
REFERENCES.....	66
APPENDIX A.....	69
MATLAB CODES.....	69
A.1 Using Neural Network:	69
A.2 Using SVR:.....	71
A.3 Using Bayesian Regression:	75

Table of Figures:

Figure 1: Comparison of load and temperature during winter and summer.	20
Figure 2: Summer profile	22
Figure 3: Winter profile	22
Figure 4: MAPE values forecasted by ANN and SD [10]	24
Figure 5: Flow chart followed by Xin and his fellows [11].	25
Figure 6: Classification Example.	27
Figure 7: Simple linear regression analysis with one independent variable.	29
Figure 8: Nonlinear regression plot for weight vs. time	31
Figure 9: Probability of passing exam to the hours of studying.	32
Figure 10: An interpolation illustration showing a finite set of points on an Epitrochoid. Spline ends in red; and interpolated curve in blue.	34
Figure 11: Given the data points in red, value at blue box, at $x=7$ is to be extrapolated.	35
Figure 12: General Flow chart describing this research.	36
Figure 13: Computation of result from the three methods.	37
Figure 14: Brief steps to be followed while using NN and SVR.	39
Figure 15: Brief steps while using Bayesian.	40
Figure 16: Image of excel database.	42
Figure 17: Brain cell structure.	43
Figure 18: Structure of Neural Network.	44
Figure 19: Simple algorithm behind SVR.	46
Figure 20: Initial command window of C compiler	47
Figure 21: K-fold cross validation method for $k=4$	49
Figure 22: a) Radial Basis Function B) RBF Mapping.....	50
Figure 23: Neural Network progress report.	52
Figure 24: Neural network model.	53
Figure 25: Plot of comparison between the trained, tested, validated and the best values.	53
Figure 26: Validation of NN model with gradient and mu values.	54
Figure 27: Histogram of the NN model.	54
Figure 28: Correlation of training, validation and testing data with the targets.	55

Figure 29: Plot between the original and the predicted values of the training data.	56
Figure 30: Plot between the original and the predicted values of the testing data.....	56
Figure 31: Original vs. predicted values of the training data with linear kernel.	57
Figure 32: Original vs. predicted values of the testing data with linear kernel.	58
Figure 33: Original vs. predicted values of the training data with RBF kernel.	59
Figure 34: Original vs. predicted values of the testing data with RBF kernel.	59
Figure 35: Original vs. predicted values of the training data with Bayesian regression.	60
Figure 36: Original vs. predicted values of the testing data with Bayesian regression.	61
Figure 37: Comparison of the three methods with original and predicted data.	62

ABSTRACT

For optimal power system operation, electrical generation must follow electrical load demand. The generation, transmission, and distribution utilities require some means to forecast the electrical load and its pricing so they can utilize their electrical infrastructure efficiently, securely, and economically. The short-term load and pricing forecast represents the electric load and cost forecast for a time interval of a few hours to a few days. This thesis will use three methods for forecasting: Neural networks, Support vector machines and the Bayesian regression. All these methods will use the same database obtained from the electrical company of Sydney, Australia. These regression models can be created and trained to receive historical load, price and future weather forecasts as inputs to produce a load and price forecast as its output. All the results from these methods will be recorded and compared to find which one gives us the best result with least MSE and NRMS.

Keywords: Electric load, Forecasting model, Regression, NN, SVM, Bayesian regression, MSE and NRMS.