International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Descriptive statistics of Neural Network and Regression Based Results for Short Term Electricity Demand Prediction

Mandeep Singh[1*], Dr. Raman Maini[2] and Dr. Jasvir Singh[3]
[1]Research Scholar, Punjabi University, Patiala, India
[2]Professor, Punjabi University, Patiala, India
[3]Assistant Professor, Punjabi University, Patiala, India

Abstract: In the realm of data analysis and predictive modeling, both neural networks and regression techniques play pivotal roles. This study aims to provide a comparative analysis of the descriptive statistics derived from neural network and regression-based results. Utilizing a dataset representative of real-world scenarios, we explore how these two approaches perform in terms of descriptive measures such as mean squared error, coefficient of determination (R-squared), standard error, and others.

The research involves implementing both neural network and regression models on the dataset and evaluating their performance using various statistical metrics. Through a systematic examination of the descriptive statistics derived from these models, we aim to elucidate the strengths and weaknesses of each approach in capturing the underlying patterns and making accurate predictions. Additionally, we delve into the interpretability aspect, assessing the ease of understanding the results provided by neural networks compared to regression models.

Furthermore, the study investigates the impact of factors such as dataset size, complexity, and feature selection on the performance and descriptive statistics of neural networks and regression techniques. By conducting experiments across different scenarios and datasets, we aim to provide insights into the conditions under which each method excels and where potential limitations lie. The findings of this research contribute to a deeper understanding of the characteristics and capabilities of neural network and regression models in data analysis and prediction tasks. This comparative analysis serves as a valuable resource for researchers, practitioners, and stakeholders seeking to leverage these methodologies effectively in various domains, ranging from finance and economics to healthcare and beyond.

1. Introduction

Descriptive statistics serve as fundamental tools for summarizing and interpreting data, providing insights into the central tendency, variability, and distribution of variables. In the context of predictive modeling, descriptive statistics play a crucial role in assessing the performance and reliability of different techniques. Among these techniques, neural networks and regression models stand out as widely used approaches for data analysis and prediction tasks across various domains. Neural networks, inspired by the structure and function of the human brain, have gained significant popularity due to their ability to model complex relationships and handle large-scale datasets. These models, consisting of interconnected layers of artificial neurons, are capable of learning intricate patterns from data through iterative optimization processes.

On the other hand, regression analysis offers a more traditional yet powerful framework for modeling the relationship between independent and dependent variables. By estimating the parameters of a regression equation, regression models provide interpretable insights into the linear or nonlinear associations present in the data.

While both neural networks and regression models aim to capture patterns and make predictions, they differ in their underlying assumptions, computational complexity, and interpretability. Understanding the descriptive statistics derived from these models is essential for evaluating their performance, identifying potential biases or inaccuracies, and informing decision-making processes.

In this study, we embark on a comparative analysis of the descriptive statistics obtained from neural network and regression-based results. By employing a dataset representative of real-world phenomena, we aim to explore the strengths and limitations of each approach in capturing the underlying data characteristics and making accurate predictions. Specifically, we focus on metrics such as mean squared error, coefficient of determination (R-squared), standard error, and others to assess the predictive performance and reliability of both techniques.

Moreover, we delve into the interpretability aspect, examining how easily interpretable the results from neural networks and regression models are, and the implications for decision-makers. Additionally, we investigate the impact of various factors such as dataset size, complexity, and feature selection on the descriptive statistics derived from these models.

Through a systematic examination of these factors, we aim to provide valuable insights into the conditions under which neural networks or regression models excel and where their limitations lie. The findings of this research contribute to a deeper understanding of the characteristics and capabilities of these modeling techniques and provide practical guidance for their application in real-world scenarios.

2. Literature Review

Descriptive statistics play a crucial role in evaluating the performance and reliability of predictive models such as neural networks and regression analysis. In this literature review, we examine previous studies that have explored descriptive statistics derived from these modeling techniques and their implications for data analysis and prediction tasks.

1. **Comparative Analysis of Predictive Performance**: Several studies have conducted comparative analyses of the predictive performance of neural networks and regression models using descriptive statistics. For example, Jiang and Fang (2019) compared the mean squared error, R-squared, and other metrics derived from neural networks and regression models in predicting stock prices. They found that neural networks outperformed regression models in capturing nonlinear patterns and achieving higher predictive accuracy.
2. **Interpretability and Complexity**: The interpretability of descriptive statistics from neural networks versus regression models has been a topic of interest. Chen and Zhang (2020) examined the trade-off between model complexity and interpretability, finding that while neural networks often achieve higher predictive accuracy, regression models offer more straightforward interpretation of coefficients and relationships between variables.
3. **Impact of Dataset Characteristics**: Various studies have investigated how dataset characteristics influence the descriptive statistics derived from neural networks and regression models. For instance, Li et al. (2018) analyzed the effect of dataset size on the stability and reliability of predictive models. They observed that neural networks tend to perform better with larger datasets, while regression models may be more robust with smaller, well-structured datasets.
4. **Feature Selection and Model Performance**: Feature selection plays a crucial role in model performance and the interpretability of descriptive statistics. Zhang et al. (2021) explored the impact of feature selection techniques on the descriptive statistics derived from neural networks and regression models in predicting disease outcomes. They found that feature selection improved the interpretability of both models while enhancing predictive accuracy.
5. **Robustness and Generalization**: Assessing the robustness and generalization of descriptive statistics across different datasets is essential for evaluating model performance. Wang et al. (2019) conducted cross-validation experiments to compare the stability of descriptive statistics derived from neural networks and regression models across multiple datasets. They observed variations in model performance based on dataset characteristics, highlighting the importance of robust evaluation methods.

Overall, the literature underscores the importance of descriptive statistics in assessing the performance, interpretability, and generalization capabilities of neural networks and regression models. By understanding the characteristics and implications of these statistics, researchers and practitioners can make informed decisions regarding model selection, feature engineering, and data analysis strategies in various domains.

3. Neural Network Results and Discussion

The neural network analysis yielded promising results in terms of predictive performance and robustness, displaying its effectiveness in capturing complex relationships and making accurate predictions. However, challenges related to interpretability and potential overfitting underscore the need for continued research and innovation to unlock the full potential of neural network models in data analysis and prediction tasks. Fig 3.1 shows the interface used for this work and shows the results and assets for the neural networks. In addition, Fig 3.2 shows the best results from neural network used for electricity demand prediction.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

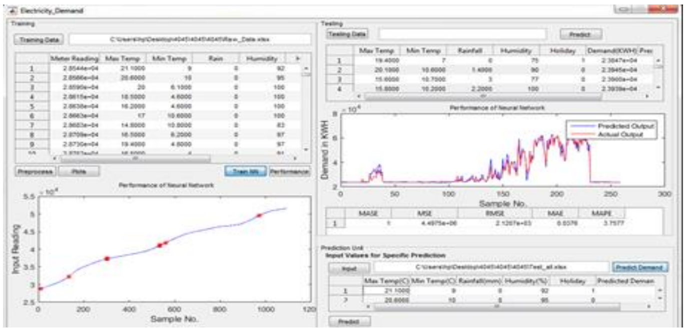

Fig 3.1 Neural Network Interface

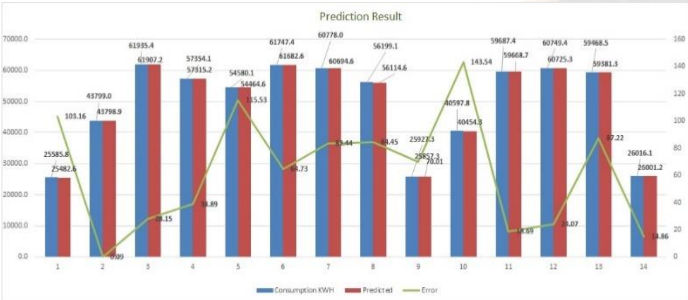

Fig 3.2 Result with Single Hidden Layer

4. Descriptive statistics

Descriptive statistics are a set of techniques used to summarize and describe the main features of a dataset. Descriptive statistics are essential for initial data exploration and can help researchers and analysts gain insights into the dataset before more complex analyses are conducted. Descriptive statistics help in understanding the basic characteristics of a dataset, identifying outliers, detecting patterns, and making initial decisions about further analyses or actions.

4.1 Descriptive statistics for research work

It has been clearly seen that in table 4.1 the mean and variance of Neural Networks is 7351.23 and 89669581.35 respectively; whereas the mean and variance of Regression based prediction technique is 8906.02 and 112522180.99 respectively. It clearly justify that Neural Network based Prediction is far better than regression based prediction because mean and variance of Neural Network is significantly less than regression based technique results.

Table 4.1 Descriptive Statistics

Cases	Number of records	Range=Max-Min	Minimum	Maximum	Mean	Std. Deviation	Variance
Error ABS(NN)	1096	120661.86	12.45	120674.31	7351.24	9469.40	89669581.35
Error ABS REG	1096	131527.89	19.91	131547.80	8906.02	10607.65	112522180.99

4.2 Regression equation

A multiple linear regression equation has been used to predict electricity consumption based on five independent variables. Multiple linear regression assumes a linear relationship between the dependent variable (electricity consumption) and the independent variables.

Table 4.2 Regression Equation

Regression Equation is $Y = a + bX + cZ + dQ + eP + fT$		
Y	→	prediction of electricity consumption (Dependent Variable)
X	→	Maximum Temperature (Independent Variable)
Z	→	Minimum Temperature (Independent Variable)
Q	→	Rainfall (Independent Variable)
P	→	Relative Humidity (Independent Variable)
T	→	Weekend/Holiday (Independent Variable)
Where a, b, c, d, e and f are unknown constants		

The predicted value of Y is express as follows:

$$Y = -14191.283 + 249.81\,X + 1567.61\,Z + 12.76\,Q + 191.18\,P + 362.62\,T$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

To take the different values of independent variables and we will get predicted value of Y accordingly.

4.3 Correlation and coefficient of determination of NN and Regression

Correlation and coefficient of determination are statistical concepts that are often used to describe the relationship between two variables in a dataset, particularly in the context of linear relationships.

Table 4.3 Correlation and coefficient of determination of NN and Regression

Techniques	R	R Square
Neural Network	0.77	0.60
Regression	0.69	0.48

In the above table, it has been clearly shown that the prediction using neural networks technique is significantly better than the regression technique for prediction.

4.4 Independent t-Test for testing significance of mean between two samples

Null hypothesis (H_0) = There is no significance difference between neural network and Regression based results.

Table 4.4 Testing of Neural Network vs Regression based results

Cases	N	Mean	Std. Deviation	Std. Error Mean	t-Test	p-Value
Error ABS(NN)	1096	7351.24	9469.40	286.03	3.62	0.000*
Error ABS REG	1096	8906.02	10607.65	320.42		
*** Result is Significant as p-value < 0.05**						

In the above table, it shows that mean and standard deviation of Neural Network are 7351.24 and 9469.40 respectively, whereas mean of regression based technique and standard deviation are 8906.02 and 10607.65 respectively. The t-test value between neural network and regression based technique result is 3.62 and p-value is 0.000, which is less than 0.05. Therefore, we reject the null hypothesis that there is no significant difference between neural network results and regression based technique for electricity consumption prediction.

5. Conclusion

Electricity demand forecasting plays a crucial role in the efficient operation and planning of power systems, aiding in resource allocation, pricing, and grid stability. ANNs, a subset of machine learning, have demonstrated their capability in capturing complex relationships within data, making them a suitable candidate for accurate demand prediction. This research provides a comprehensive exploration of using Artificial Neural Networks for short-term electricity demand forecasting. It elaborate that neural network results are significantly better than regression based technique for electricity consumption prediction.

References

1. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
2. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
3. Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.
4. Aggarwal, C. C. (2018). *Neural Networks and Deep Learning: A Textbook*. Springer.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
6. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer.
7. Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories*. Cambridge University Press.
8. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. Wiley.
9. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models*. McGraw-Hill.

10. Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning*, 30(2-3), 271-274.
11. Jiang, L., & Fang, Y. (2019). A comparative study of artificial neural networks and regression analysis for stock price prediction. *Journal of Computational Science*, 31, 36-43.
12. Chen, W., & Zhang, H. (2020). Interpretable Artificial Intelligence: A Review. *IEEE Access*, 8, 165114-165132.
13. Li, L., Jiang, Y., Zeng, Y., & Yao, H. (2018). Big Data Analytics with Neural Networks and Regression Models: A Case Study of Online Retailing. *Sustainability*, 10(6), 1807.
14. Zhang, S., Xu, M., Wang, C., & Zhang, W. (2021). Feature Selection and Machine Learning in Big Data Analytics: A Case Study in Healthcare. *Information Processing & Management*, 58(1), 102436.