Masters in data science

Tools and Techniques in Data Science Project Report

Submission Date: 20-10-2024

Group Members:

Ganesha Maheshwari (DS-035/2024)

Muhammad Saleem Siddiqui (GD-002/2024)

Hamza Hanif Alam (GD-001/2024)

Submitted to: Dr. Erum Abbasi

Introduction:

The rapid evolution of power systems, driven by the integration of renewable energy sources, increasing demand, and growing grid complexity, has emphasized the need for advanced state estimation techniques. Conventional state estimation methods, while effective in many scenarios, face significant challenges when dealing with the dynamic and large-scale nature of modern power distribution systems. To address these limitations, our project, **"Neural Network-Based Distribution System State Estimation"**, introduces an innovative approach leveraging the power of neural networks.

This cutting-edge method aims to enhance real-time monitoring and operational decision-making in complex distribution networks by offering faster and more accurate state estimation. By utilizing machine learning techniques, specifically neural networks, the proposed solution can effectively manage large distribution systems' nonlinearity and scalability challenges. The ultimate objective is to improve grid reliability, operational efficiency, and adaptability, which are essential for the transition to smarter and more sustainable energy systems.

Our project sets a foundation for modernizing state estimation practices, paving the way for intelligent and responsive power grids capable of meeting future demands.

Background:

In modern power systems, state estimation is a fundamental process used to determine the system's operating conditions by estimating the electrical states, such as voltages and phase angles, across the network. This capability is crucial for grid monitoring, stability analysis, and real-time operational control. Traditional state estimation methods, such as Weighted Least Squares (WLS), rely on mathematical models and measurement data to approximate system states. These methods have been effective in many applications but face significant limitations when applied to large and complex power distribution networks.

One of the major challenges with conventional techniques is the high computational time required to process large volumes of data in expansive networks. This delay can hinder real-time monitoring and responsiveness. Additionally, accuracy issues may arise due to model assumptions, incomplete measurement data, and non-linearities inherent in distribution systems. These factors limit the ability of traditional methods to meet the demands of modern power grids, which are becoming increasingly dynamic due to renewable energy integration, distributed generation, and fluctuating loads.

To address these challenges, recent advancements in artificial intelligence (AI) and machine learning (ML) have paved the way for innovative approaches to state estimation. Neural networks, a subset of ML, have shown exceptional capabilities in handling complex, non-linear problems and processing large datasets with high efficiency. By leveraging these strengths, neural network-based methods can provide faster, more accurate, and scalable solutions for state estimation.

Our project builds on this foundation, aiming to develop a neural network-based state estimation framework tailored to the needs of modern power distribution systems. This approach not only addresses the limitations of traditional methods but also aligns with the ongoing evolution of smart grids, enabling enhanced real-time monitoring, improved reliability, and greater operational efficiency.

Methodology:

The methodology for our project, **"Neural Network-Based Distribution System State Estimation,"** is structured into the following key steps to ensure a comprehensive and robust implementation:

1. Test System Selection

We selected the **IEEE 6-bus radial distribution system** as the test case for our study. This system is widely used in power system research due to its ability to represent typical distribution network characteristics while maintaining simplicity for analysis and experimentation.

2. Data Generation

To generate the data required for training and testing the neural network models, we implemented a **load flow algorithm**. The **backward-forward sweep method** was chosen because it is specifically designed for radial distribution systems and provides computational efficiency with reliable convergence.

- Input data includes load profiles, system topology, line parameters, and other network characteristics.
- Output data consists of system states, such as bus voltages and line currents, under various loading conditions.

3. Incorporating Realistic Noise and Errors

To simulate real-world operational conditions, we introduced **realistic noise and measurement errors** into the generated dataset. This step ensures that the neural network models are trained to handle data imperfections, thereby enhancing their robustness and generalization capabilities for practical applications.

4. Neural Network Design and Training

We developed a neural network-based framework for state estimation with the following design considerations:

- **Architecture**: A multilayer feedforward neural network (MLP) with sufficient hidden layers and neurons to capture the non-linear relationships in the data.
- **Inputs**: Network measurements (e.g., power flows, injected powers) and system configuration data.
- **Outputs**: Estimated state variables, such as bus voltages.
- **Training**: The neural network is trained using the generated dataset, optimized with techniques such as backpropagation and stochastic gradient descent.
- **Validation**: The model's performance is validated using a portion of the data reserved for testing, ensuring it generalizes well to unseen scenarios.

5. Performance Evaluation

The trained neural network is evaluated based on its:

- **Accuracy**: Comparison of estimated states with ground-truth values from the load flow algorithm.
- **Computational Efficiency**: Assessment of processing time compared to conventional methods.
- **Robustness**: Ability to maintain performance under noisy and imperfect data conditions.

6. Iteration and Optimization

If performance metrics indicate areas for improvement, the model architecture, training parameters, or noise assumptions are iteratively refined to achieve better results.

7. Result Analysis and Interpretation

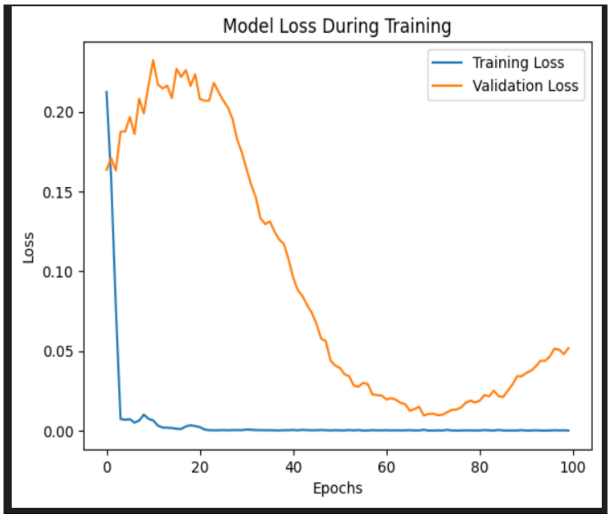
The final step involves analyzing the outcomes to assess the feasibility and scalability of the proposed approach for larger and more complex distribution networks. The insights gained will guide further advancements and potential real-world implementation.

This systematic methodology ensures that the neural network-based state estimation framework is both reliable and applicable to modern power distribution systems.

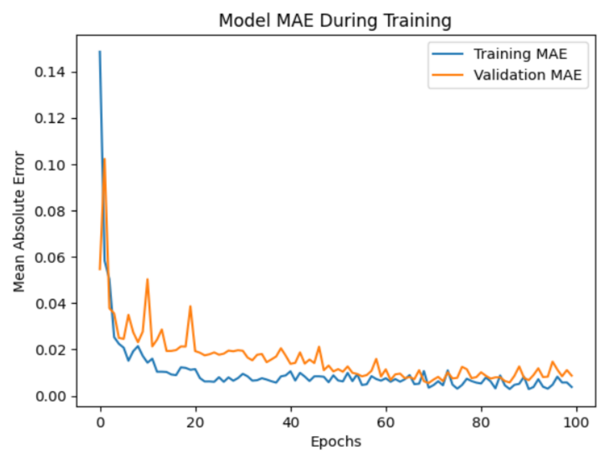
Results of Predictive Models

Model 1 Graphs:

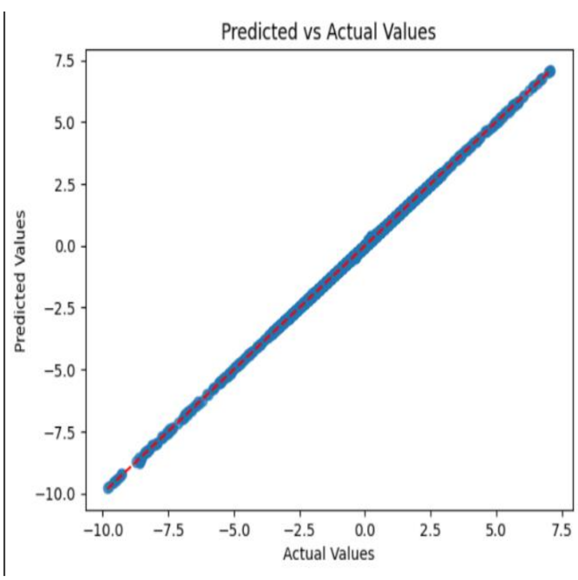
Visualize Training History



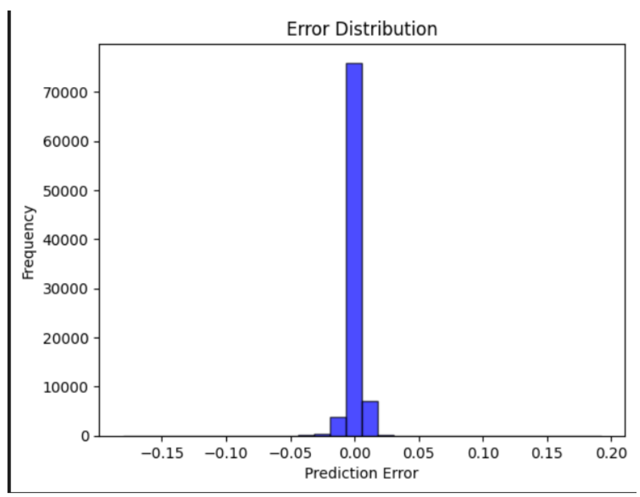
Training and Validation MAE vs Epochs



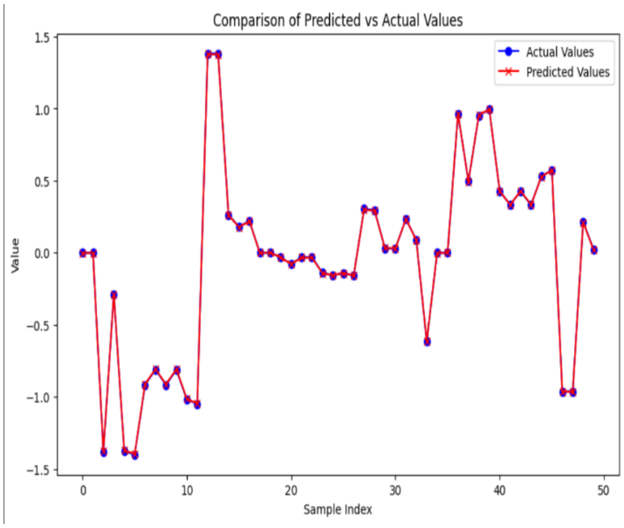
Predicted vs Actual Values (Scatter Plot)

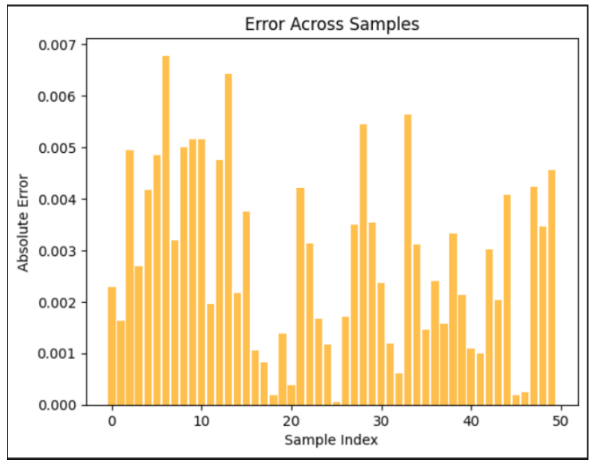


Error Distribution (Histogram)

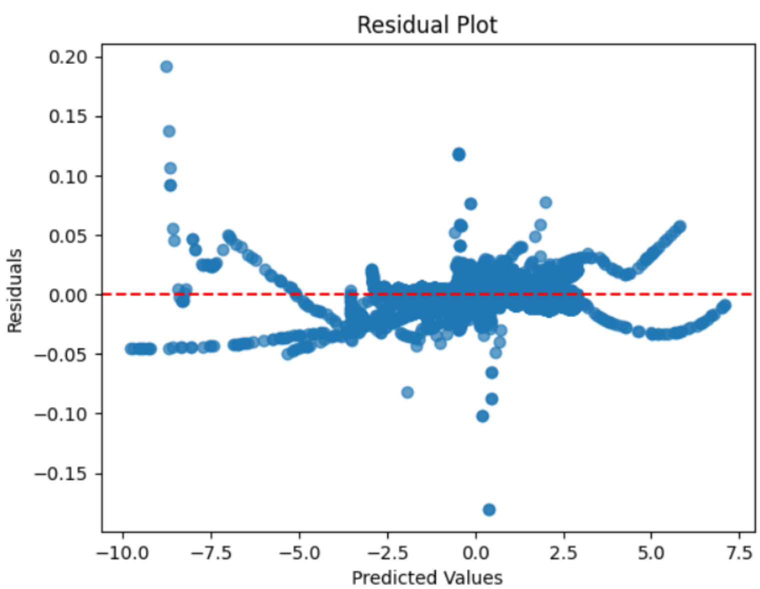


Comparing of Actual and predictive value **Error Across Samples (Bar Plot)**



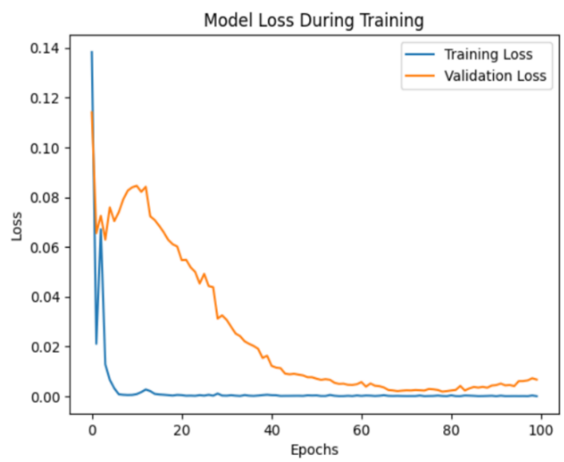


Residual Plot

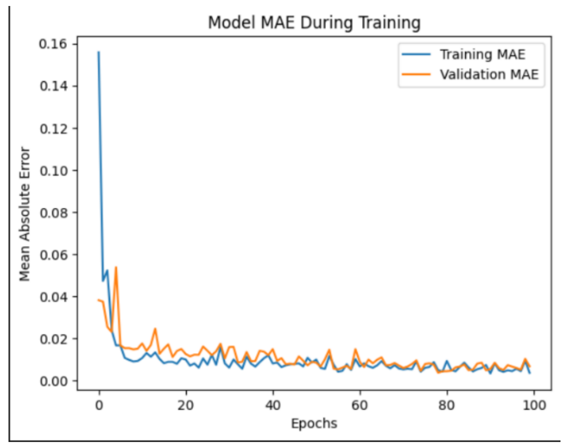


Model 2 Graphs:

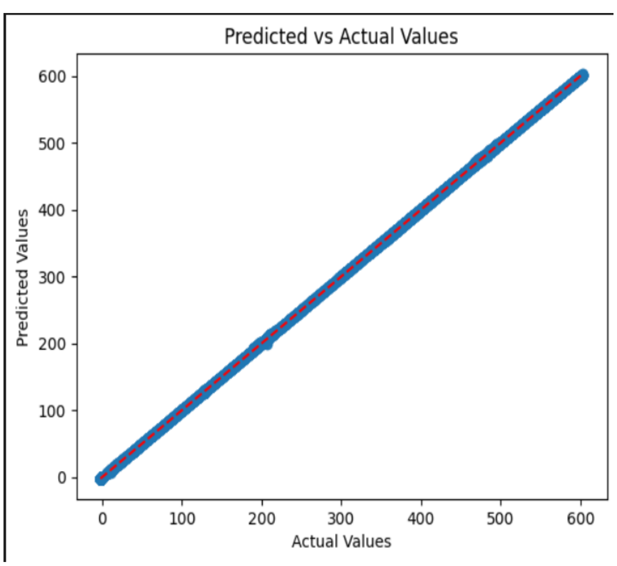
Visualize Training History



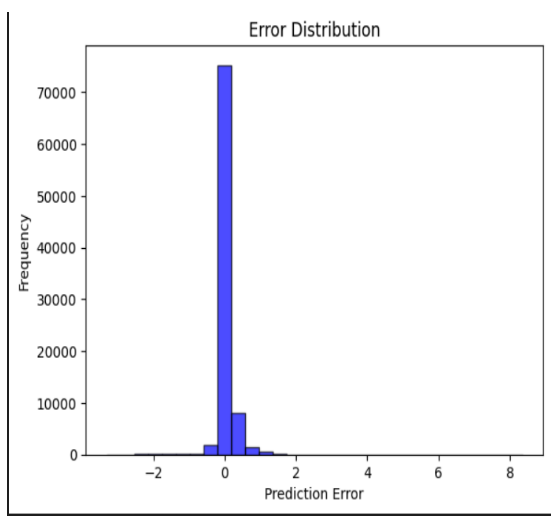
Training and Validation MAE vs Epochs



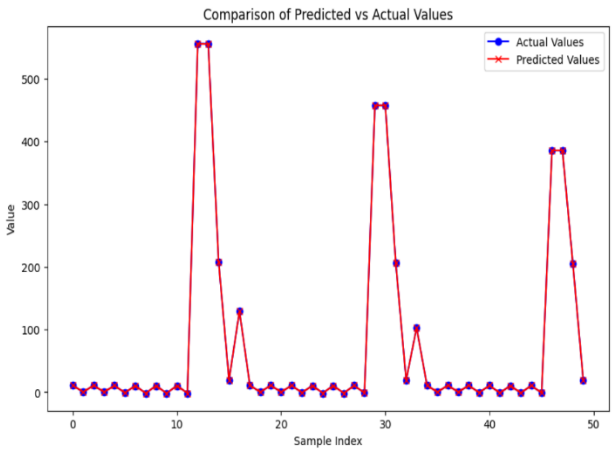
Predicted vs Actual Values (Scatter Plot)



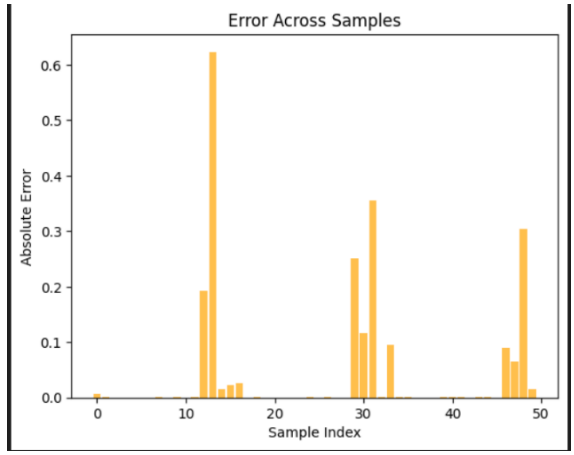
Error Distribution (Histogram)



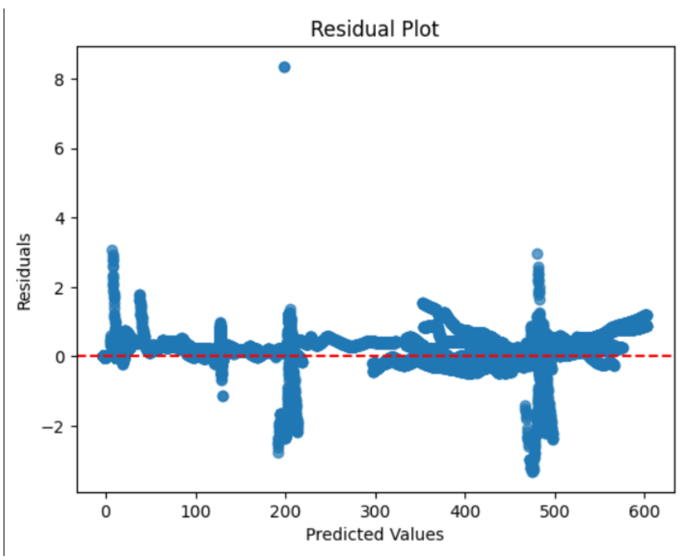
Comparing of Actual and predictive value



Error Across Samples (Bar Plot)



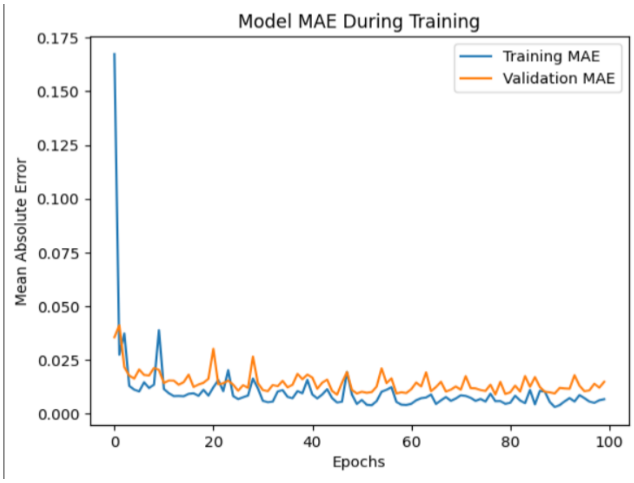
Residual Plot



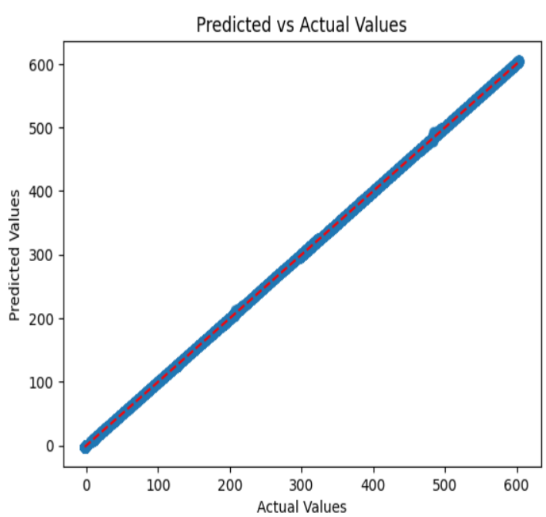
Model 3 Graphs:

Visualize Training History

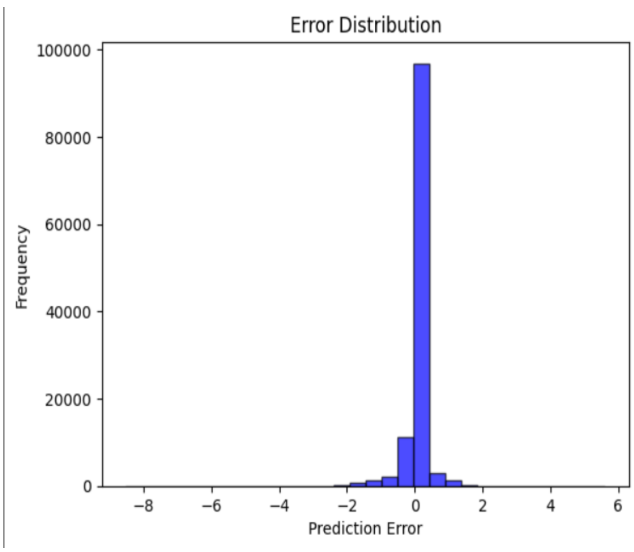
Training and Validation MAE vs Epochs



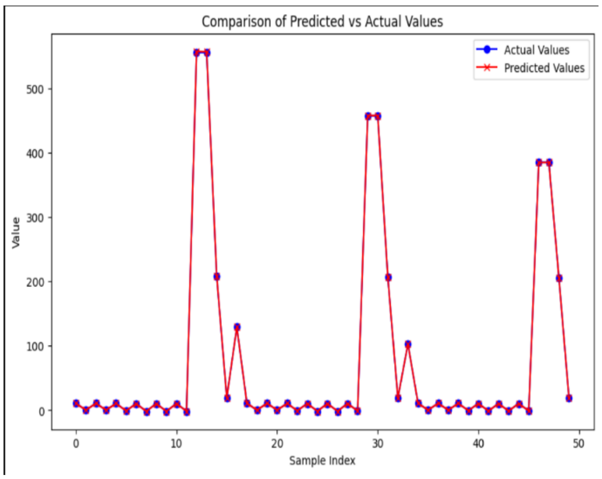
Predicted vs Actual Values (Scatter Plot)



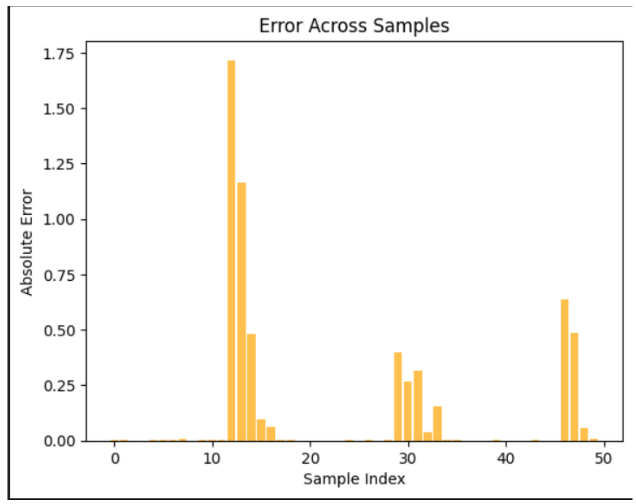
Error Distribution (Histogram)



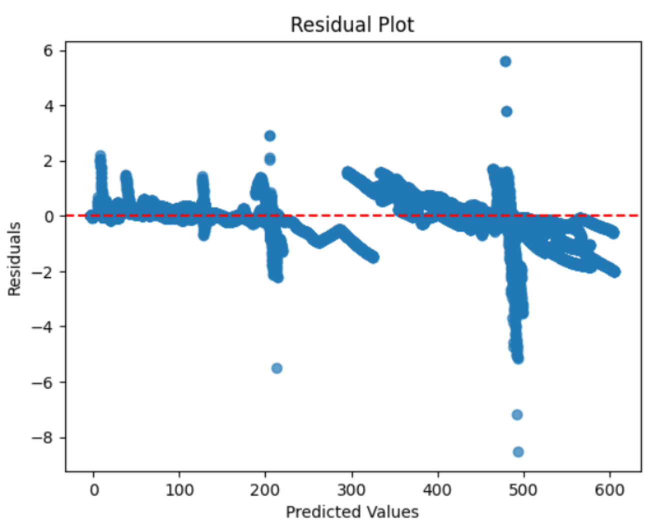
Comparing of Actual and predictive value



Error Across Samples (Bar Plot)



Residual Plot



Conclusion

Our project has successfully completed the initial phase of data generation, creating a comprehensive dataset tailored for neural network-based state estimation in power distribution systems. Building on this dataset, we developed and evaluated artificial neural network (ANN) models using three different training and testing configurations.

Model Training and Testing Configurations

1. **Case 1:** Training with 80% of the dataset and testing with 20%.
2. **Case 2:** Training with 70% of the dataset and testing with 30%.
3. **Case 3:** Training with 60% of the dataset and testing with 40%.

The results from these configurations demonstrate that the performance of the neural network improves as the proportion of training data increases. In Case 1, where 80% of the data was used for training, the model achieved the lowest error rates. As the training data reduced to 70% and 60% in subsequent cases, the error rates increased, reflecting the impact of reduced training data on the network's learning ability.

Key Observations

- **Error Reduction:** The model achieved the highest accuracy in Case 1 (80% training), where error rates were significantly reduced due to the larger training dataset.
- **Testing Generalization:** While increasing the testing proportion (as in Cases 2 and 3) provides a broader evaluation of the model, it also highlights the trade-off between training data size and model performance.

This phased approach to training and testing validates the robustness of the ANN models and underscores the importance of sufficient training data for accurate state estimation. These results signify a promising step toward developing ANN-based solutions that can revolutionize the speed and accuracy of power distribution system monitoring. By addressing computational challenges and enhancing real-time capabilities, our work contributes to the foundation of more efficient and reliable grid operations.