Predicting Power Load Using Deep Learning

Assigned by	Prof. Jiehan Zhou
	jiehan.zhou@ieee.org
Completed by	
Ayub Ali Emon	202486060005@sdust.edu.cn
Phoutthavong Phonesamay	202486060004@sdust.edu.cn

Shandong University of Science and Technology

ABSTRACT

In this study, we employed a Long Short-Term Memory (LSTM) model to predict the energy load data based on historical energy usage and meteorological data. The load data was collected from Arizona State University's Tempe campus and meteorological data from the National Renewable Energy Laboratory. Our goal was to accurately forecast energy consumption, leveraging time-series modeling techniques. We have chosen LSTM to address the challenge of fluctuating energy loads in large-scale facilities.

After several model iterations, hyperparameter tuning, and feature selection, we achieved a final Mean Absolute Percentage Error (MAPE) of 1.99% and a Root Mean Squared Error (RMSE) of 9606.14 KW. These results demonstrate the model's strength in learning patterns from the input features and predicting future energy loads with minimal error. Despite challenges with initial flat predictions, adjustments in feature engineering, and sequence length we have significantly improved the model's performance.

This report outlines the methodology, model development process, challenges encountered, and final results, providing insights into using LSTM models for time-series prediction in energy management.

All the data related to this study process is available in the following GitHub repository:

github.com/alfa-echo-niner-ait/loadlstm

Contents

1. Introduction

Energy consumption forecasting plays a crucial role in the management and optimization of energy systems, especially in large-scale buildings. Accurately predicting energy loads helps in efficient energy distribution, reducing carbon emissions, and optimizing operational costs.

The objective of this study is to develop a Long Short-Term Memory (LSTM) based model that can predict the energy load data. The prediction is based on historical energy consumption data and meteorological information, such as temperature and humidity. The aim is to explore how well LSTM, as a time-series forecasting method, can capture the temporal dependencies in energy consumption data and provide accurate predictions.

This report outlines the steps involved in building and fine-tuning the LSTM model, from data preprocessing and model selection to performance evaluation and analysis. The final model demonstrates strong predictive power, providing a useful tool for energy load forecasting in real-world applications.

2. Data Processing

In any machine learning project, data preprocessing is a critical step that directly affects the performance of the model. In this study, the data used for energy load prediction was collected from two main sources:

- **Energy load data** from the IES building at Arizona State University's Tempe campus, which includes information such as total energy consumption (KW) and greenhouse gas emissions (GHG).
- **Meteorological data** sourced from the National Renewable Energy Laboratory provides key weather-related variables such as temperature, humidity, and wind speed, all of which impact the energy usage of a large building.

The collected data spanned two months (from 05/01/2024 to 06/30/2024) and contained daily readings, resulting in a dataset with 61 entries and 36 features, including variables such as KW, CHWTON, HTmmBTU, temperature, humidity, wind speed, and station pressure.

2.1 Data Cleaning and Merging

Before training the model, it was essential to clean and preprocess the data to ensure its quality. Several steps were followed:

- **Handling missing values**: Any missing values in the dataset were either filled with the mean of the respective feature or interpolated based on neighboring values. In cases where a feature had a significant amount of missing data, it was excluded from the analysis.

- **Outlier detection**: Outliers in the energy load and meteorological data were identified using statistical methods (e.g., z-scores) and were either capped or removed based on their impact on the overall distribution.
- **Merging datasets**: The energy load data and the meteorological data were merged on the date field, ensuring the alignment of both datasets for daily observations. The merging process resulted in a unified dataset that could be used for training the LSTM model.

2.2 Feature Selection and Normalization

Selecting the right features is crucial for the model's performance. The following features were selected for normalization based on their relevance to energy consumption:

Feature Column	Description	Example
KW	The total kilowatt (KW) power consumption for the day.	351973.45 KW
GHG	Greenhouse gas emissions.	169.754
AVG Air Temperature [deg F]	The average air temperature in degrees Fahrenheit for the day.	76.2268 def F
AVG Rel Humidity [%]	The average relative humidity as a percentage for the day.	19.3349%
AVG Wind Speed @ 3m [MPH]	The average wind speed at 3 meters above ground level in miles per hour for the day.	7.5972 mph

These features were chosen based on their impact on energy load fluctuations, as environmental factors like temperature, humidity, and wind speed significantly affect a building's heating, cooling, and ventilation requirements. From here we select the KW feature as our target column.

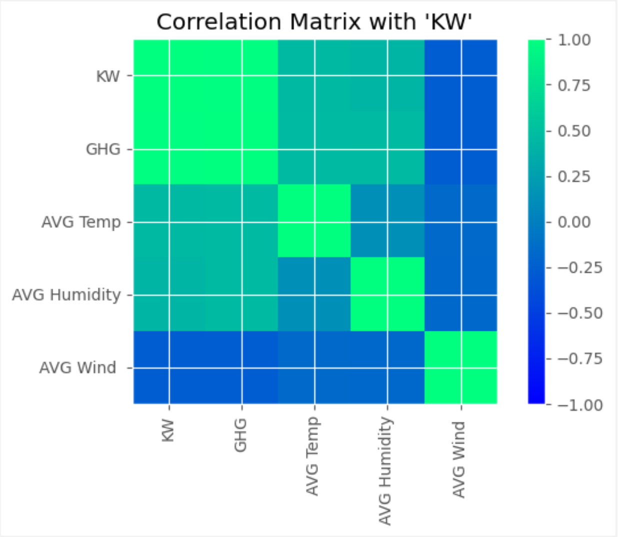

Fig 2.1: Selected Features Correlation with Target Column KW

The correlation matrix shown in Fig 2.1 insights into the relationships between the energy load (KW) and various environmental factors such as GHG emissions, average temperature, humidity, and wind speed. There is a moderately strong positive correlation between KW and GHG, indicating that higher energy loads are associated with increased greenhouse gas emissions. A weaker positive correlation is observed between KW and average temperature, suggesting that energy load tends to rise slightly with higher temperatures, likely due to increased cooling demand. Conversely, KW and average wind speed exhibit a strong negative correlation, implying that higher wind speeds correlate with reduced energy load, possibly due to cooler conditions or alternative wind-based energy sources. The relationship between KW and humidity appears nearly neutral, showing little direct influence of humidity on energy consumption. These correlations are important for understanding how environmental factors impact energy load, potentially improving model predictions.

After selecting the features, normalization was performed to scale the data. Normalization is particularly important for LSTM models because it helps stabilize the learning process and ensures that the gradient descent algorithm converges more effectively. In this study, the *Min-Max Scaling* method was applied to transform the features into a range of [0, 1]. This step ensured that no single feature dominated the others during training, and it also improved model convergence.

3. Methodology

This section outlines the steps and decisions taken to develop an accurate and efficient Long Short-Term Memory (LSTM) model for predicting the energy load. The methodology is divided into four key components: model selection, feature engineering, model architecture, and training process.

3.1 Model Selection

Given the time-series nature of the energy load data, the LSTM model was selected as the primary model for this work.

Why LSTM?

- **Sequential Data Handling**: LSTM networks are well-suited for time-series prediction tasks because of their ability to retain long-term dependencies. Unlike standard recurrent neural networks (RNNs), LSTMs are designed to mitigate the vanishing gradient problem, which allows them to "remember" patterns over longer sequences of data.

- **Energy Load Forecasting**: Energy usage typically follows a daily, weekly, or seasonal pattern influenced by both human activities and weather conditions. The ability of LSTMs to handle such temporal dependencies made them a strong candidate for this application.

- **Previous Studies**: Many previous studies in the energy sector have shown that LSTM-based models outperform traditional machine learning models such as ARIMA or decision trees when it comes to energy consumption forecasting.

3.2 Feature Engineering

Feature engineering is crucial to maximize the predictive power of the LSTM model. The features selected for the model, as described earlier, included both energy load and environmental metrics. These features were selected because they have a direct impact on building energy usage.

- **KW (Kilowatts)**: Represents the total energy load and is the target variable for the prediction.
- **GHG (Greenhouse Gas Emissions)**: An indicator of carbon emissions associated with energy usage, serving as an indirect measure of load.
- **AVG Air Temperature [deg F]**: Outdoor temperature, which influences cooling and heating loads in the building.
- **AVG Rel Humidity [%]**: Humidity levels, which affect cooling requirements.
- **AVG Avg Wind Speed @ 3m [MPH]**: Wind speed, which can influence heating and cooling efficiency.

After selecting the relevant features, normalization was applied to ensure that all features contributed equally during training. Min-Max scaling was used, transforming each feature to the [0,1] range.

3.3 Model Architecture

The LSTM model was designed with the following architecture:

- **Input Dimension**: The model's input consists of 5 features (KW, GHG, temperature, humidity, and wind speed), representing the normalized time-series data.
- **Hidden Layers**: The model uses 4 LSTM layers, with each layer consisting of 64 hidden units (or dimensions). The choice of 64 hidden dimensions was the result of experimentation, balancing model complexity and training performance.
- **Dropout**: To prevent overfitting, a dropout layer with a 10% dropout rate was added between each LSTM layer. This regularization technique randomly drops units during training to improve generalization.
- **Output Layer**: A single dense (fully connected) layer was used to output the final predicted energy load value.

3.4 Training Process

The model was trained using backpropagation through time (BPTT) and gradient descent techniques. The following configurations and techniques were employed during the training process:

- **Loss Function**: Mean Squared Error (MSE) was chosen as the loss function since it penalizes large prediction errors, making it appropriate for a regression task like energy load forecasting.
- **Optimizer**: Adam optimizer was used with a learning rate of 0.0005, incorporating L2 regularization (weight decay) with a factor of 0.001 to reduce overfitting. Adam's adaptive learning rate mechanism made it suitable for handling the stochastic nature of time-series data.
- **Scheduler**: The learning rate scheduler (ReduceLROnPlateau) was applied to dynamically reduce the learning rate by a factor of 0.1 if the model's performance on the validation set did not improve for 20 consecutive epochs.
- **Sequence Length**: The model used a sequence length of 8 days of past data to predict the energy load for the next day. This choice was based on several trial-and-error analyses, balancing the model's ability to capture temporal patterns and the risk of overfitting.
- **Batch Size and Epochs**: The model was trained for 80 epochs with a batch size of 32, which was selected to optimize the balance between training speed and model convergence. Gradient clipping with a maximum norm of 7 was applied to prevent exploding gradients during training.

Key Hyperparameters:

Learning rate	0.0005
Weight decay (L2 regularization)	0.001
Batch size	32
Number of epochs	80
Gradient clipping: Max norm	7

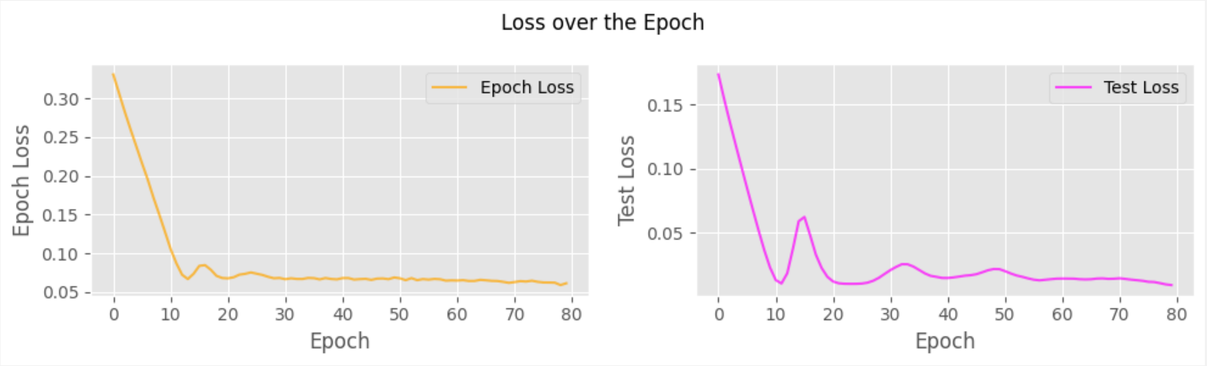

Fig 3.1: Training Process

The training process, as shown in *Fig 3.1*, indicates a successful convergence of the model. The training loss rapidly decreases during the first 10 epochs, stabilizing around 0.05 after that, suggesting that the model learns effectively without much fluctuation. However, the test loss exhibits more variability, particularly between epochs 10 and 30, with noticeable spikes around epochs 10 and 20. This may indicate some overfitting or challenges in generalizing during this period. After epoch 30, the test loss stabilizes, following a similar trend as the training loss, but with minor fluctuations. Overall, the model appears to have converged well, with both losses reaching low values.

4. Model Evaluation and Results

In this section, we present the evaluation of the LSTM model's performance based on the metrics calculated on the test data. The evaluation process involved both initial experiments that provided insights into the model's behavior and a final performance assessment, including comparisons between predicted and actual energy load values.

4.1 Initial Experiments

During the initial experiments, the model faced several challenges that impacted its performance:

- **Flat Predictions**: The model exhibited flat predictions for several epochs, indicating potential overfitting. This behavior was addressed through various regularization techniques, including dropout layers and L2 regularization.

- **Tuning Sequence Length**: Different sequence lengths were tested to determine the optimal amount of historical data to include. An initial choice of 7 days did not yield satisfactory results, leading to the selection of an 8-day sequence, which significantly improved performance.

- **Early Stopping**: The introduction of early stopping was initially considered to prevent overfitting; however, the results indicated that training the model for the full 80 epochs provided better performance.

Through iterative tuning and experimentation, the model's configuration was optimized to achieve satisfactory performance.

4.2 Final Model Performance

The final model was evaluated on a test dataset, yielding the following performance metrics:

- **Mean Absolute Percentage Error (MAPE)**: 1.99%

- **Root Mean Square Error (RMSE)**: 9606.14 KW

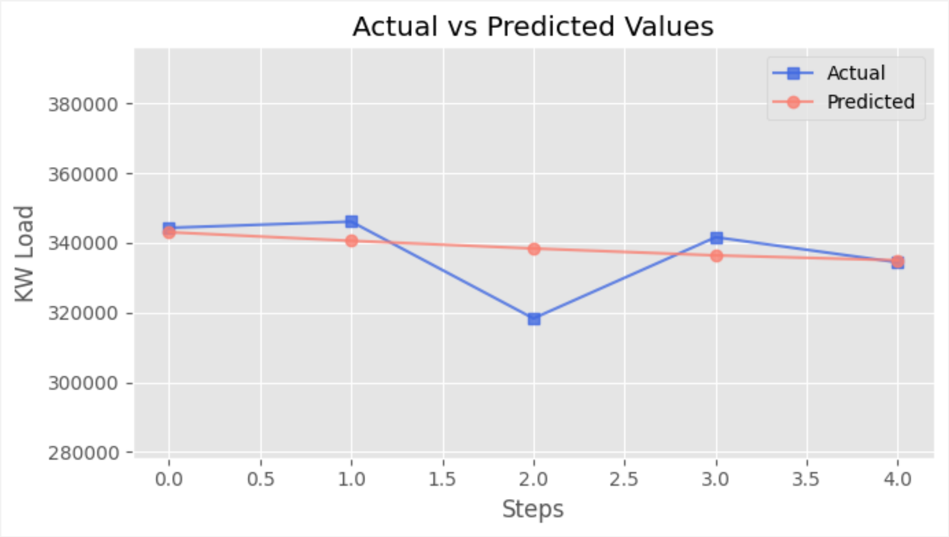

Fig 4.1: Model Testing

Fig 4.1 shows the comparison between actual and predicted KW load values across five steps. The actual values, represented by the royal blue line, display more variability, particularly around step 2, where the KW load dips sharply below 320,000. The predicted values, represented by the salmon line, follow a much smoother trajectory, maintaining closer to 340,000 KW across all steps. While the model predicts well for steps 0, 1, 3, and 4, where the actual and predicted values are closely aligned, there is a noticeable divergence at step 2, where the actual value drops significantly while the predicted value remains relatively stable. This suggests that while the model performs well overall, it struggles to capture sudden changes in the KW load, leading to some underestimation during periods of high variance.

4.3 Analysis of Errors

The model evaluation revealed several key insights regarding the prediction errors:

- **Error Distribution**: The largest discrepancy occurred for the third predicted value, where the model underestimated the actual energy load by over 20,000 KW. This suggests that external factors or anomalies not captured in the model could have influenced energy consumption on that particular day.
- **Factors Influencing Accuracy**: Various factors, including weather conditions, occupancy levels, and operational characteristics of the IES building, can significantly impact energy consumption. These factors might introduce variability that the model struggles to predict accurately.
- **General Trend**: Despite some errors, the model successfully captured general trends in energy consumption, suggesting it can be effectively utilized for short-term forecasting in energy management applications.

5. Challenges and Lessons

The process of developing and fine-tuning the LSTM model for energy load prediction involved several challenges, each offering valuable insights that contributed to the final results. This section discusses the obstacles encountered and the lessons learned throughout the modeling process.

5.1 Challenges Faced

- **Difficulties in Model Training**: During training, the model exhibited several challenges, including flat predictions during initial epochs. This behavior often indicated that the model was either overfitting to the training data or not sufficiently capturing the underlying patterns in the energy load data.

- **Tuning the Sequence Length**: Selecting the optimal sequence length was a critical aspect of model performance. An initial choice of a shorter sequence led to less accurate predictions. Through experimentation, an 8-day sequence length was determined to provide the best results, demonstrating the importance of historical context in time-series predictions.

- **Experimenting with Mechanisms**: The team explored the possibility of incorporating an attention mechanism to enhance the model's focus on relevant time steps. However, the added complexity did not yield a significant performance improvement compared to the standard LSTM configuration. This led to the decision to revert to the simpler model, emphasizing the need for a balance between model complexity and interpretability.

5.2 Lessons Learned

- **Importance of Feature Selection**: Careful selection of features proved to be a pivotal factor in model performance. The features chosen for normalization—KW, GHG, average air temperature, relative humidity, and average wind speed—were critical in capturing the relevant dynamics influencing energy consumption.

- **Significance of Hyperparameter Tuning**: The tuning of hyperparameters, such as learning rate, dropout rate, and batch size, had a considerable impact on the model's accuracy. The iterative process of experimenting with different configurations highlighted the necessity of rigorous hyperparameter optimization to achieve the best results.

- **Insights into LSTM Behavior**: The study provided insights into how LSTM models function in the context of energy load forecasting. Specifically, it reinforced the understanding that LSTMs can effectively capture temporal dependencies when provided with adequate historical data and appropriate feature representations.

- **Adaptability to Real-World Data**: The challenges encountered during modeling demonstrated the adaptability required when working with real-world data, which often contains noise and variability. This adaptability is crucial for developing robust predictive models in energy management applications.

6. Conclusion

The development of an LSTM-based model for predicting the energy load represents a significant step towards leveraging machine learning techniques in energy management. This study successfully integrated energy consumption data with meteorological factors, enabling the model to capture complex temporal patterns that influence energy demand.

Summary of Key Findings

- The final LSTM model demonstrated strong performance metrics, achieving a Mean Absolute Percentage Error (MAPE) of 1.99% and a Root Mean Square Error (RMSE) of 9606.14 KW. These results indicate that the model is capable of providing accurate predictions for future energy loads based on historical data.
- The decision to utilize a sequence length of 8 days allowed the model to effectively leverage historical patterns, improving predictive accuracy compared to shorter sequences.
- The study highlighted the importance of feature selection, hyperparameter tuning, and model complexity. While the exploration of an attention mechanism was insightful, the final model configuration reaffirmed that a simpler architecture can yield superior results in certain contexts.

Implications for Real-world Energy Management Applications

The findings from this study have significant implications for real-world energy management. By accurately predicting energy loads, stakeholders can optimize energy usage, reduce operational costs, and contribute to sustainability efforts. This model can serve as a foundational tool for facility managers and energy analysts, enabling them to make informed decisions based on reliable forecasts.

6.1 Future Work

Moving forward, there are several potential avenues for improvement and exploration:

1. **Incorporating Additional Features**: Future iterations of the model could benefit from integrating additional data sources, such as occupancy levels or appliance usage, to enhance prediction accuracy.
2. **Exploring Other Machine Learning Techniques**: Investigating other advanced machine learning approaches, such as ensemble methods or hybrid models combining LSTM with other architectures, may yield further improvements in predictive performance.
3. **Longer Time Horizon Predictions**: Expanding the model's capabilities to predict energy loads over longer time horizons could provide valuable insights for strategic planning and resource allocation.

4. **Real-Time Predictions**: Implementing the model for real-time energy load predictions could enable dynamic energy management solutions, adapting to changing conditions and usage patterns.

In summary, this study not only contributes to the understanding of LSTM applications in energy load forecasting but also sets the stage for future advancements in predictive modeling within the field of energy management. The combination of accurate modeling techniques and real-time data utilization can significantly enhance energy efficiency and sustainability efforts in large buildings and beyond.

Acknowledgement

We would like to express our sincere gratitude to **Prof. Jiehan Zhou** for his invaluable guidance and support. We also extend our appreciation to **Arizona State University** for providing the energy load data from their **Tempe campus**, and the **National Renewable Energy Laboratory** for the meteorological data that made this study possible. In addition, we utilized the help of **open-source GPT** models and valuable resources from the internet to further enhance our understanding of model training and hyperparameter tuning. This contributed significantly to achieving our final model results.