Enhancing Urban Environments: The Smart Street Lighting Project using LSTM-DNN Hybrid Model

Yeni Jung

Student ID: 2308283

BSc Artificial Intelligence and Computer Science

40 Credit Project

Supervised by: Qamar Natsheh

8700 words

School of Computer Science

College of Engineering and Physical Sciences

University of Birmingham

2023-24

Abstract

This research project focuses on develop artificial intelligence to improve urban environments by optimising street lighting energy consumption using climate data. The goal is to implement a hybrid artificial intelligence model that utilises datasets on streetlight energy usage and weather conditions to predict future energy demands based on anticipated weather patterns. By exploiting the predictive capabilities of individual AI models concerning both streetlight energy consumption and local weather conditions, this study aims to develop a comprehensive framework capable of accurately forecasting near-future energy requirements. This approach is designed to enhance urban planning and energy efficiency, thereby contributing significantly to the development of smart, sustainable urban lighting systems.

Acknowledgements

I would like to express my gratitude to my supervisor, Qamar Natsheh, for the unwavering academic and personal support during my final year project. I am also grateful to my parents and peer friends for their mental assistance that helped achieving the successful completion of my BSc.

Abbreviations

AI	Artificial Intelligence
LSTM	Long Short Time Memory
DNN	Deep Neural Network
RNN	Recursive Neural Network
CNN	Convolutional Neural Network
GBM	Gradient Boosting Model

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

In the 21st century, urban environments have become the central point for technological advancements and environmental issues. In cities with more than half the world's population, cities represent complex ecosystems where human activities, economic transactions, and ecological impacts intersect. These intersections make the urban centre a hub of innovation and a focus for sustainable development. [4]

Urban planners and environmentalists are increasingly facing challenges such as high energy consumption, rising pollution levels, and the overall impact of climate change. These problems threaten the quality of life, economic stability, and environmental health of urban areas. Addressing these challenges requires innovative approaches that transcend traditional methods and focus on sustainable, long-term solutions. [41]

Among these challenges, artificial intelligence (AI) and data analysis have emerged as innovative tools in urban planning. Artificial intelligence applications significantly demonstrate the potential to improve efficiency and sustainability, from optimising traffic flow and public transportation to managing utilities and resource distribution. By utilising artificial intelligence, cities can become more innovative and more responsive to the needs of their residents and environment. [55]

Climate data is a critical factor in this technological evolution. Detailed and accurate climate data is essential to understanding the environmental impact of urbanisation and planning interventions that mitigate side effects while promoting sustainability. Integrating AI and climate data analysis presents a promising future for urban planners. It provides the potential to anticipate and address environmental issues in advance and respond to them. [18]

The project explores how AI can be used for climate data analysis to optimise urban environments and maximise energy efficiency simultaneously. Focusing on synthesising AI and climate data, the study seeks to find patterns and insights that lead to more effective and sustainable urban planning decisions.

1.1 Research Motivation

The increasing pressure on energy resources and the environmental impact of urban areas emphasises the urgent need for sustainable urban development. As cities expand, energy consumption also increases, often outpacing the growth of sustainable energy solutions. This imbalance highlights the critical challenge of improving energy efficiency in urban environments. [14] Furthermore, the effects of climate change, such as rising temperatures and unpredictable weather patterns, worsen the challenges facing cities, and it is essential to optimise urban environments for better sustainability. [24]

Artificial intelligence offers a new approach to addressing these challenges by enabling the analysis of large-scale climate data. This data, which covers a wide range of variables from temperature fluctuations to precipitation patterns, is highly effective in understanding and mitigating the impact of urbanisation on the environment. AI's ability to process and analyse massive datasets quickly and accurately can identify trends and anomalies that traditional methods cannot. These insights will enable proactive resource management and strategic planning rather than reactive responses that fail to address the underlying issues fully. [55, 48]

Integrating AI with climate data analytics can also help develop intelligent, adaptive urban infrastructure that responds dynamically to human and environmental changes. For example, AI can optimise street lighting systems in real time to reduce energy waste and adapt to the actual needs of urban areas based on environmental data inputs. This twin focus on optimisation and energy efficiency helps reduce operational costs and contributes significantly to achieving broader ecological sustainability goals. [54]

Efforts have been made to link limited energy usage with weather to maximise energy efficiency through the potential of such artificial intelligence. [10] However, there has never been such maximisation for urban and regional units, such as street lighting energy consumption. This research project aims to fill this gap by systematically exploring how AI can harness climate data to improve the urban environment. In doing so, it is expected to provide actionable and scaleable ideas that can be implemented in urban planning and management to achieve the twin goals of improving environmental quality and energy efficiency.

1.2 Aims and Objectives

This project aims to design and implement an integrated artificial intelligence model that utilises streetlight energy usage and weather conditions datasets to predict future energy demand based on expected weather patterns. By harnessing the capabilities of individual artificial intelligence models to predict specific variables, namely the energy consumption of streetlights and local weather conditions, the project aims to develop a comprehensive framework that can accurately predict energy demand in the next 12 hours. In this way, the urban environment can be optimised and achieve maximised energy efficiency, enabling smart street lighting systems to become essential to sustainable urban development.

To accomplish these goals, various research objectives were considered:

- Collect and organise street light usage data and climate data
- Preprocess the datasets and perform Exploratory Data Analysis (EDA)

- Develop and optimise machine learning models for each data type for individual prediction.
- Integrate individual models to form an advanced energy usage prediction.
- Test the model with testing data and evaluate the accuracy of the integrated model.

1.3 Thesis Overview

Figure 1.1 shows a brief outline of the proposed methodology. In the first step, streetlight energy usage and weather datasets are collected and preprocessed through data cleaning and exploratory data analysis. Next, these preprocessed datasets are used as training and evaluation data to train an LSTM predictive model for predicting each dataset, and the model's hyperparameters are tuned through Bayesian optimisation. The optimal LSTM models obtained through these processes are added as layers to the DNN, which is an integrated model. Bayesian optimisation is performed on the obtained LSTM-DNN model to get the optimised parameters, which are then used to retrain the integrated model. The model is finally evaluated using several metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2).

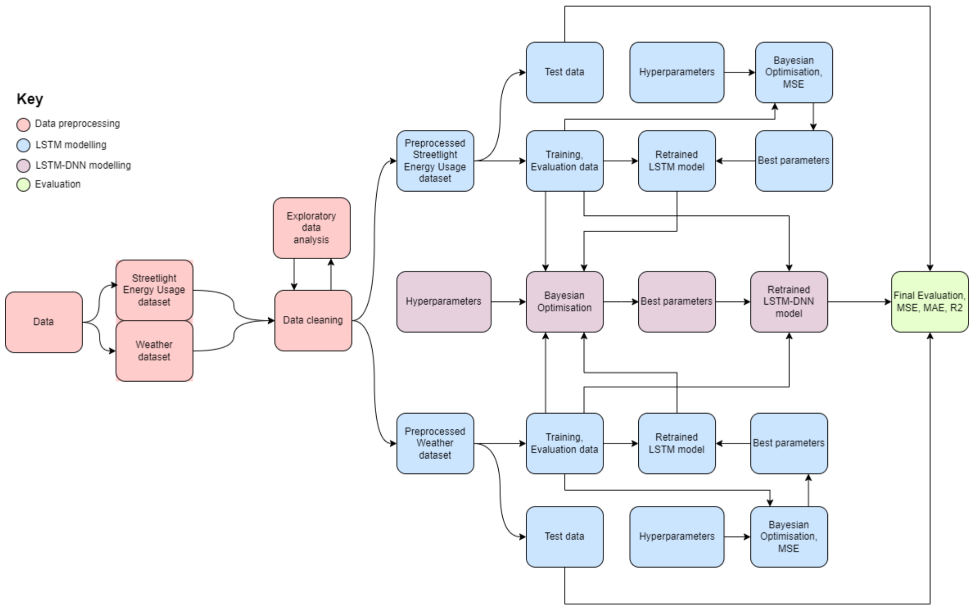

Figure 1.1: The primary principle of the proposed strategy

1.4 Limitations

The suggested system has essential limitations that should be acknowledged. Firstly, as the datasets are collected from the actual world, they are noisy and are considered an effect of external factors. Since taking this into account is challenging, and there are restrictions on the technology, the following proposed system will not propose complex methods of dealing with the noise. Secondly, the street lighting energy usage dataset used in this study is based on traditional street lighting systems that are primarily active only at night and, therefore, do not appear to consume any energy outside the nighttime hours set by the street lighting control system. This data configuration has the limitation that it does not reflect the variability in energy use during the day. For example, special events or dramatic climate changes during daylight hours can affect energy use patterns, but this dataset does not allow such situations to be analysed.

Despite these limitations, the predictive model developed in this study still provides useful value. By predicting the energy consumption patterns of streetlights, the model can improve energy management efficiency in cities. Furthermore, this study suggests to future researchers the need to develop methodologies that compensate for the existing data limitations and allow for a more precise analysis of streetlight energy use under different environmental conditions. Therefore, with additional data collection, including daytime data, the model has the potential to evolve into a more sophisticated and comprehensive energy management tool. Advancements in research such as this will facilitate the development of smart city technologies and play a essential role in energy usage prevention. Specific directions for future work, which will elaborate on this predictive model's potential enhancements and methodological advancements, will be discussed in more detail in section 5.2.

CHAPTER 2

Background

This chapter provides the reader with some required technical knowledge to understand the models and formulas utilised in this study. If the reader needs further understanding, the book "Introduction to artificial intelligence and neural networks" by V. Sangeetha and S. Andrews is recommended. [42]

2.1 Feature Scaling

Feature scaling is the act of equalising the range of values for different variables to prevent bias in model training, depending on the size of the data range. This study used one technique.

2.1.1 Normalisation

Normalisation transforms features into a different range, usually [-1,1] or [0,1]. This study used scikitlearn's MinMaxScaler to scale the data between [0,1]. The equation for this process is [56]:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

2.2 Forecast Models

2.2.1 LSTM

LSTM is a type of RNN designed to address the long-term dependencies of RNNs. It predicts future data by considering the immediate past and, more macroscopically, past data. The core idea of an LSTM is a cell state where information can only be explicitly written or removed, ensuring the state remains constant without external interference. This state can only be modified by specific gates, which are ways of passing information through. A typical LSTM comprises three gates: the forget gate, which determines whether the cell state will be deleted; the input gate, which determines what new information will be added to the memory cell; and the output gate, which determines the total cell output. [47] The LSTM model used in this project was implemented using *Keras* library (version 3.2.1).

2.2.2 DNN

A deep neural network (DNN) is an artificial neural network that contains multiple hidden layers, which is used to learn complex properties and patterns of input data. The network consists of an input layer, several hidden layers, and an output layer, and is utilized in various fields such as speech recognition, image recognition, and natural language processing. DNNs have a powerful ability to learn complex nonlinear relationships. [20] The DNN model used in this project was implemented using *Keras* library (version 3.2.1).

2.3 Evaluation Metrics

2.3.1 Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAE is calculated by converting the difference between the actual value y_i and predicted value $\hat{y}_i$ into absolute values and sum them to get the average. [17]

2.3.2 Mean Squared Error

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

MSE is calculated by squaring the difference between the actual value y_i and the predicted value $\hat{y}_i$ and averaging. [17]

2.3.3 Coefficient of Determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

The coefficient of determination is also known as R2 or R-squared. It represents the percentages of variance between the predicted result $\hat{y}_i$ and the actual value y_i, and the mean $\bar{y}_i$ with the value range of [0,1]. The closer the value gets to 1 the better the model performs. The negative value indicates the model is evaluated incorrectly. [17]

CHAPTER 3

Literature Review

This literature review will bridge the sophisticated critical literature of available energy consumption prediction models while encompassing the broad spectrum of models on both sides, offering extensive primary data collection and statistical background knowledge. It, therefore, puts the previous progress and remaining issues into context, allowing for the innovative features of the latest studies to be thoroughly discussed. The more specific details of the models included highlight the issue-based value of the debate: the interaction between many key data input factors above, including the interestingly compelling advent of factoring in both LED and weather data to enhance the quality of the predictions. The upcoming sections will review this literature in a critical analysis of the overall developments of total models from conventional to the most sophisticated-lite variants of AI.

3.1 Traditional Model

In traditional modelling approaches, techniques such as artificial neural networks [45], random forest algorithms, and support vector machines [51] have been utilised to predict energy consumption. However, although these methods can be mining features, they have only sometimes provided optimal predictive performance. The limitations often stem from challenges in selecting and extracting relevant features - a crucial step that remains complex for energy prediction tasks. [43, 32] Deep learning and machine learning techniques have expanded upon these traditional methods. Studies have leveraged these advanced techniques for more accurate predictions of energy consumption. Mohan et al. [37] developed a dynamic-empirical energy prediction model, deemed unreliable due to its high computational demands. In contrast, statistical learning models applying autoregressive models such as ARIMA [7] and GARCH [5] offer a less resource-intensive solution. Khodayer et al.[30] employed the GARCH model to address predictive uncertainty, yet it struggled to handle the nonlinearity and nonstationarity of energy consumption data. Research incorporating linear regression has also been conducted, with Fumo et al. [19] developing a model for residential energy consumption prediction. Vu et al. [49] and Braun [8] extended this work with multi-regression models. Nonetheless, it was found that statistical energy consumption models often need to capture uncertainty trends more directly; alternative methods are instead used to reduce prediction errors. [43] Chen et al. [12] utilised support vector regression to forecast seasonal energy consumption, integrating the SVR model with the fruit fly optimisation algorithm to enhance performance. Cao and Wu [9] applied a hybrid SVR model for a similar purpose, while Zhong et al. [57] took a different approach using a vector field-based SVR model for building energy consumption. Li et al. [34] addressed the nonlinearity challenge in the energy consumption dataset through a random forest regression algorithm combined with fast Fourier transforms for frequency prediction.

3.2 Deep Learning Models

Several models were proposed in various works of literature to predict energy consumption. In the following literature, it is suggested that Kong et al. [33] provide a model of LSTM for energy consumption prediction. The research forecasted the energy consumption model by rolling update and attention algorithm with the help of the bidirectional LSTM model [50]. Similarly, Shi et al. proposed a pooling deep RNN model for the prediction of household energy, highlighting the challenge of overfitting within some feature extraction algorithms [43]. The feature extraction techniques consist of several regular patterns and analysis of spectrum and noise. On the contrary, methods reduce the accuracy level of energy consumption, whereas algorithms lower the accuracy due to non-linearity and irregular trends of the dataset of electrical people consumption. Deep learning techniques, particularly LSTM [33] and Convolutional Neural Networks (CNN) [40], have been recognised for their proficiency in mastering the trend of patterns and sequences. However, CNN models often struggle to capture temporal features within energy consumption data, leading to the proposal of Gated Recurrent Unit (GRU) models by Han et al. [22] for predicting future energy usage within short intervals, thereby reducing error rates. Ullah et al. [46] utilised a hybrid bidirectional LSTM and CNN model to overcome the limitations of singular algorithm models, while Kim et al. [31] combined the CNN-LSTM model for predicting energy and household consumption. Chi et al. [13] introduced the WT multiple LSTM model to enhance the accuracy of power consumption prediction. Khan et al. [29] developed the CNN-BiGRU model, reporting the lowest forecasting error in electricity consumption for individual dwellings. Furthermore, they proposed the CNN-ESN model [28], which detailed accuracy in consumption forecasting without demanding high operational bandwidth. Alsharekh and collaborators [2] advanced this field with an ML-RNN model, a deep R-CNN model equipped with enumerated attention mechanisms, showcasing significant achievements in forecasting short-term electricity load under actual database experiments. Expanding upon these methodologies, recent advancements have introduced hybrid models and deep learning approaches, emphasising the increasing accuracy and computational efficiency of energy consumption predictions. A notable development is the integration of Deep Neural Networks (DNN) with LSTM units, forming a hybrid DNN-LSTM model that surpasses traditional models in accuracy, particularly by effectively handling complex, non-linear patterns in energy usage data [1]. This model's success is echoed in exploring deep learning models for building energy consumption prediction, where a deep learning-based model demonstrated superior performance compared to machine learning and statistical regression-based benchmark models, especially in predicting the cooling energy consumption of buildings [3]. These advancements enhance predictive accuracy and contribute to developing more sophisticated, accurate, and user-friendly forecasting methodologies in energy consumption analysis.

3.3 Climate Influence on Energy Consumption

These studies have illustrated the significant role of considering how weather conditions can be used to improve energy consumption. For example, in the relationship between temperature and electricity consumption, the study postulated the existence of a U-shaped model, indicating a non-linear association. These represent the more pronounced effects with apparent implications for energy demand, especially for urban and industrial areas [35]. Through this, the news does well to marry with the study in pushing forward knowledge in the area by considering how smart city infrastructure might use data, such as that associated with weather information, to deliver more energy-efficient use. This special issue seeks to collect work explaining the importance of weather data integration in energy planning and policy formulation, all geared toward reducing energy impact in cities through consumption optimisation strategies considering climatic conditions. [38] These studies have combined to illustrate how weather conditions affect energy usage. The primary considerations in this relationship may include temperature, air pressure, dew point, wind speed, and others. Weather metrics, yet again, are visible in the list of sophisticated data analysis methods for energy data analytics, coupled with predictive models, to present the commanded precision to leverage an energy platform for future improvement in distribution and utilisation efficiency for sustainable urban energy. Recent work has brought revolutionary developments in deep learning architecture models for predicting energy consumption, considering the impact of weather and atmospheric variables. Chung and Jang's recent paper [15] further reiterates the thesis proposed in this article: climatic data will require being grouped with another piece of information for improved knowledge. Applying the CNN-LSTM model has shown that deep learning models have a more excellent capability than traditional machine learning models. In the study of Mehmood et al., [36], adjustable parameters such as air pressure, dew point, and wind speed were integrated into the optimisation procedure, showing that it can result in significant energy saving without compromising user comfort. The inclusion of weather data has revealed an emerging trend in energy management strategies, aiming to exploit the subtle climate effect on energy demand in prediction models. These works have been saliently brought out in effectiveness within the energy sector to influence consumption patterns. This reflects a positive and efficient management practice toward energy by improving the accuracy of demand forecasts through complex, sophisticated, and comprehensive embedded prediction models that incorporate a wide range of weather metrics.

3.4 Knowledge Gap

Recent research on prediction models for energy consumption explored various possibilities. It used deep learning frameworks, such as LSTM and CNN, and combinations of them to find different energy usage patterns. An essential aspect of this research is integrating LEDs and weather data to provide a more reliable energy consumption forecast. At the same time, all of these studies were focused on forecasting at local scales, for example, within buildings or other property boundaries. Such an approach, particularly concerning the use of LEDs, can underestimate or overestimate the real influence on consumption.

Moreover, the methodological strategy of using separate AI models to predict weather and energy usage, with subsequent complex and detailed connections, has yet to be explored. The proposed project methodology suggests integrating individual LSTM models to analyse streetlight energy consumption and weather datasets. This will be followed by integration through DNN. This approach aims to address the identified knowledge gap by providing the ability to predict public energy consumption at the urban level beyond the limited property unit. Moreover, the implementation and integration of two distinct deep learning architectures ensure the accurate capture of temporal dependencies and patterns specific to each dataset. This is an important aspect that many traditional models often need to pay more attention to, as they may need to account for time-series nuances or rely on generalised patterns. The subsequent unification of these analyses through a DNN can leverage its capability to identify complex, non-linear relationships between the nuanced patterns detected by the LSTM models. This two-step methodology involves a separate detailed analysis followed by an integrated synthesis, which allows for a depth of understanding and predictive accuracy that conventional single-model approaches may not achieve. Such approaches may only partially exploit the relations between diverse data types. Moreover, this approach offers scalability and adaptability, which allows the model to integrate supplementary datasets or adjust to new predictive challenges effortlessly. This confers a notable advantage over more rigid or narrowly focused predictive models. The potential impact of this approach on energy policy planning, urban development, and sustainability efforts is substantial, as it provides more accurate data for more sustainable energy use.

CHAPTER 4

Methodology

4.1 Data

This chapter fully describes the data used in this study. It gives the source of the data and explains its attributes. Then, any preprocessing, including transforming the original data and dealing with missing values, is described. Methods of exploratory data analysis are elaborated further.

4.1.1 Data Collection

The "Street Light Energy Usage" dataset, obtained from the Data World platform and attributed to the City of Las Vegas, Nevada, was recorded over two years, between October 8, 2014, and October 12, 2016. [16] The data is recorded every 15 minutes, offering keen granularity depending on different times and conditions. This dataset encompasses several key attributes:

- Basic Information: This includes each street light's name and location (labelled *'loc'*) and billing and administrative information such as *'account', 'customer_code'*, and *'premise_code'*.
- Metre details include *'meter_id'*, *'service_type'*, and *'channel_number'*, specifying the metre's identity and the type of service recorded.
- Energy Usage Data: This section contains *'power_flow', 'unit'* of measurement, *'interval_length'* of recording, *'time_zone'* of data, and timestamps (*'start_date_time', 'end_date_time', 'start_date', 'end_date', 'start_time', 'end_time'*) that define the recording periods.
- Consumption and Cost Metrics: These metrics are detailed by *'time_of_use', 'read_type'* indicating manual or automatic readings, *'usage'* quantities, *'day_of_week', 'hour_of_the_day'*, and the *'cost'* associated with the energy consumed.

Figure 4.1is a map generated using the Python programming language and the *Folium* library to visualise the 15 locations of each street light in Las Vegas.

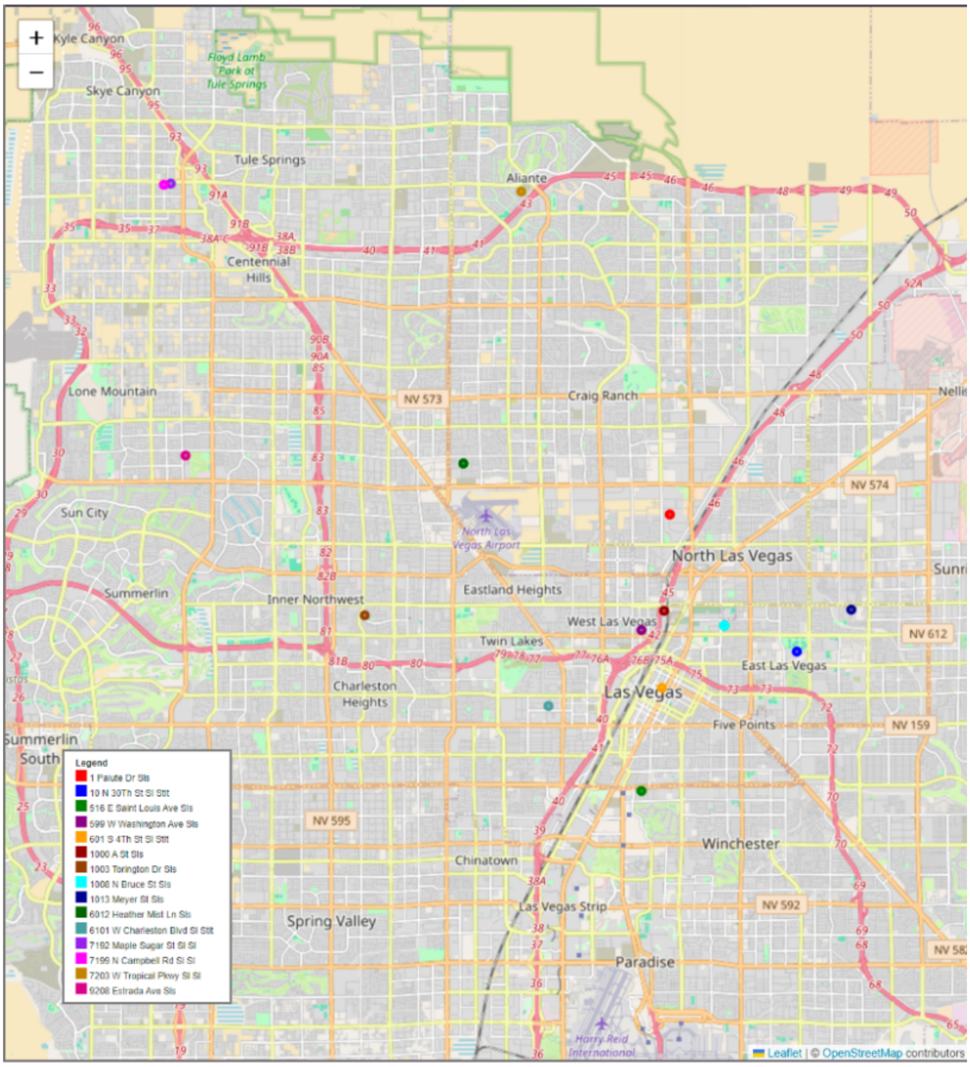

Figure 4.1: The locations of 15 different street lights of the 'Street Light Energy Usage' dataset.

Table 4.1 selectively presents the most relevant attributes of the dataset, providing a structured and visual overview of its critical elements.

The weather dataset was sourced from Weatherbit.io via their publicly available API, enabling precise acquisition of detailed weather conditions at 15-minute intervals in Las Vegas, aligning with the streetlight energy usage dataset for comprehensive analysis. The API documentation provided extensive guidance on retrieving the data, ensuring efficient collection according to the platform's terms of use. The dataset was initially obtained in JSON format through the API and was subsequently transformed into CSV format to facilitate manipulation and analysis using various data processing tools.

The projects mentioned have followed the terms and conditions set by Weatherbit.io, with a particular emphasis on using the dataset exclusively for academic and non-commercial purposes, as outlined in the

Table 4.1: Preview statistics of the "Street Light Energy Usage" dataset

Name	Loc	Unit	Start Date Time	Usage	Cost
CITY OF LAS VEGAS	1 Paiute Dr Sls	KWH_DEL	10/08/2014 12:00:00 AM	0.926	$0.00
CITY OF LAS VEGAS	1 Paiute Dr Sls	KWH_DEL	10/08/2014 12:15:00 AM	0.927	$0.00
CITY OF LAS VEGAS	1 Paiute Dr Sls	KWH_DEL	10/08/2014 12:30:00 AM	0.925	$0.00
CITY OF LAS VEGAS	1 Paiute Dr Sls	KWH_DEL	10/08/2014 12:45:00 AM	0.926	$0.00
CITY OF LAS VEGAS	1 Paiute Dr Sls	KWH_DEL	10/08/201401:00:00 AM	0.927	$0.00

guidelines stated in the Terms and Conditions Document [52], which specifies that 'Non-commercial – You may not use the Material for commercial purposes'. This description highlights the commitment to the ethical use of the dataset by established usage policies.

To meet the specific requirements of our study, particularly in comparison with the LED dataset, the weather data was adjusted to match the same observation period, geographical region, and interval timing. The data collection process involved gathering information from 15 street light metres at 15-minute intervals over two years, from October 8, 2014, to October 13, 2016. To improve the model's accuracy and capture annual seasonal weather patterns more effectively, it was decided to include data from the previous five years (8 October 2010 to 7 October 2014) in developing the individual LSTM model.

The dataset contains several attributes that are essential for weather analysis:

- Temperature-related: *'app_temp'* (apparent temperature), *'temp'* (ambient air temperature), *'dewpt'* (dew point), *'rh'* (relative humidity percentage)
- Solar-related: *'dhi'* (Diffuse Horizontal Irradiance), *'dni'* (Direct Normal Irradiance), *'ghi'* (Global Horizontal Irradiance) *'solar_rad'* (total solar radiation received by the earth), *'elev_angle'* (solar elevation angle above the horizon), *'azimuth'*(solar azimuth angle, showing the sun's position), *'uv'* (ultraviolet index, the level of solar UV radiation)
- Precipitation related: *'precip_rate'* (precipitation rate), *'snow_rate'* (rate of snowfall)
- Wind-related: *'wind_dir'* (wind direction in degrees), *'wind_gust_spd'* (speed of wind gust), *'wind_spd'* (wind speed)
- Others: *'pod'* (part of the day - day/night), *'ts'* (Unix timestamp), *'timestamp_local'* (local timestamp), *'timestamp_utc'* (coordinated universal timestamp), *'revision_status'* (data revision status)

4.1.2 Data Preprocessing

Streetlight Energy Usage Dataset

In this study, the initial streetlight energy usage dataset contained numerous parameters - 24 in total (refer to subsection 4.1.1) - of which only *'start date time'*, *'loc'* (location) and *'usage'* were considered central to the analysis as they had a direct impact on assessing energy consumption patterns. The conscious choice to exclude the remaining 21 parameters, including customer identifiers (*'name', 'account number', 'customer code', 'premise code'*) and various technical details (*'meter ID', 'service type', 'channel number'*), and retain only these, was driven by the intention to extract the dataset down to its most analytically valuable elements. This strategic pruning was done to remove features that, while potentially informative in other contexts, may contribute little to the current analysis or introduce unnecessary complexity or variability.

For example, customer and location identifiers were excluded because this study aims to understand broad patterns of energy use across multiple locations, not the behaviour of individual customers. Despite their operational relevance, technical details and operational metrics were deemed unnecessary for the study, as the *'usage'* parameter summarises essential information about energy consumption. Similarly, excluding temporal parameters such as *'start_date', 'start_time'*, etc. and using a consolidated *'start date time'* aims at a more streamlined temporal analysis. In addition, utility and cost factors such as *'hours_of_use', 'read_type', 'cost', and 'end_date_time'* have been omitted as they are less relevant to the primary goal of understanding physical energy use patterns. The guiding principle behind these selective exclusions is Occam's Razor, which favours simplicity and focuses the analysis on parameters most likely to yield meaningful insights into energy consumption trends. [23] This methodological choice not only improves the predictive model's performance by reducing noise and potential bias but also aims to improve the performance and interpretability of the model by simplifying the dataset to the essential features that influence energy consumption patterns by focusing the analysis on variables that directly affect energy use.

Initially presented as strings, the location values (labelled *'loc'*) were converted to integer values to ensure compatibility with numerical analysis techniques and machine learning algorithms. This process, detailed in Table 4.2, provides a more structured dataset, allowing efficient processing and analysis.

The dataset has been pivoted to improve visibility and facilitate efficient data handling. The DataFrame index was set as the *'start date time'*, column indices were transposed from *'loc'* values, and cells were populated with 'Usage' values, greatly simplifying the time series analysis approach. This reorganisation allows energy consumption to be reviewed chronologically and compared across different locations simultaneously. Furthermore, this arrangement simplifies input preparation by presenting each location as a distinct feature. It ensures data completeness, enabling the LSTM to effectively capture and learn from temporal patterns across multiple series.

Despite the description of the data resources, there is some variability in the measurement periods across different locations, as shown in Table 4.2. Intersecting recording periods from October 15, 2014, to October 5, 2016, were selected to promote data consistency. This section will examine the impact of this decision on the dataset's size and the representativeness of the analysis period. Temporal coverage was standardised across all locations by excluding data outside this intersection period. The review evaluates the impact of the exclusion on the findings, including potential biases that could reduce or introduce data diversity. This is important to confirm the integrity of the dataset and the reliability of subsequent analyses within a defined recording period.

Table 4.2: Consolidated Overview of Location Identifier Conversion and Measurement Period Analysis

Original location value	**Location identifier**	**Measurement range**
1 Paiute Dr Sls	1	8th Oct 2014 - 5th Oct 2016
10 N 30Th St Sl Stlt	2	8th Oct 2014 - 5th Oct 2016
516 E Saint Louis Ave Sls	3	15th Oct 2014 - 12th Oct 2016
599 W Washington Ave Sls	4	15th Oct 2014 - 12th Oct 2016
601 S 4Th St Sl Stlt	5	15th Oct 2014 - 12th Oct 2016
1000 A St Sls	6	8th Oct 2014 - 5th Oct 2016
1003 Torington Dr Sls	7	8th Oct 2014 - 5th Oct 2016
1008 N Bruce St Sls	8	14th Oct 2014 - 11th Oct 2016
1013 Meyer St Sls	9	14th Oct 2014 - 11th Oct 2016
6012 Heather Mist Ln Sls	10	15th Oct 2014 - 12th Oct 2016
6101 W Charleston Blvd Sl Stlt	11	15th Oct 2014 - 12th Oct 2016
7192 Maple Sugar St Sl Sl	12	15th Oct 2014 - 12th Oct 2016
7199 N Campbell Rd Sl Sl	13	15th Oct 2014 - 12th Oct 2016
7203 W Tropical Pkwy Sl Sl	14	15th Oct 2014 - 12th Oct 2016
9208 Estrada Ave Sls	15	15th Oct 2014 - 12th Oct 2016

Upon further investigation, it was found that data needed to be included during specific periods on 8th March 2015 and 13th March 2016, specifically between 14:00 and 14:45. The gap was due to missing time steps rather than NaN values within existing rows. To address this issue, missing time steps were created, and a time series interpolation was applied to estimate the 'Usage' value for these intervals. A detailed discussion of the implementation of this interpolation method, including its theoretical foundation and practical application, is provided in Figure 4.1.2.

For the performance of the models, the dataset was segmented into training, validation, and test datasets. The data segmentation ratio follows the 6:2:2 rule, which allocates 60% of the total data to the training dataset, 20% to the validation dataset, and 20% to the test dataset. Table 4.3 shows each DateTime range of the separated datasets.

Table 4.3: The DateTime range of the train, validation, and test dataset

	DateTime range	# of data points
Train	2014-10-08 00:00:00 ∼2015-12-22 04:30:00	41587
Validation	2015-12-22 04:45:00 ∼2016-05-14 14:00:00	13862
Test	2016-05-14 14:15:00 ∼2016-10-05 23:45:00	13863

Weather Dataset

When preparing the weather dataset for integrated analysis (2014-2016) for the study, thorough steps were taken, particularly as described in subsection 4.1.2, to ensure compatibility with the streetlight energy usage datasets aligned with the observation period from 15th October 2014 to 5th October 2016. This adjustment was critical for creating input for an integrated prediction model and analysis performance. Notably, the excluded period from 8th October 2014 to 14th October 2014 remained intact.

Instead, this data was combined with the extended dataset (2010-2014) to ensure the time consistency of the past data and the integrated data.

An initial column pruning process to remove irrelevant information from the analysis was made for the integrated analysis (2014-2016) and the extended dataset for individual prediction (2010-2014). *'Timestamp_utc'* and *'ts'* (unix timestamp) have been removed in favour of local time representations, which are crucial to match weather data with local energy consumption patterns accurately. The *'weather'* parameter, mainly consisting of textual forecast information, has been considered non-essential for quantitative analysis, which relies on numerical data for predictive modelling. Similarly, *'dhi'* and *'dni'* were consolidated into *'ghi'* based on solar energy research, which indicated that GHI is adequate for assessing solar potential. [6] Additionally, *'pod'* (part of the day) and *'slp'* (sea level pressure) were excluded due to their limited impact on the study's specific weather and energy usage prediction objectives. The dataset was carefully selected to prioritise variables directly impacting the analysis, improving model performance and interpretability.

To match the streetlight energy usage dataset, the name of the *'timestamp_local'* column was changed to *'start date time'*, ensuring the creation of the input structure for the integrated prediction model. Furthermore, examining the integrity of the dataset, duplicated data items were found from 2010 to 2015, especially during certain intervals between 01:00 and 01:45 the day after the daylight savings adjustment. These redundancies were addressed by averaging and maintaining the integrity of the dataset by ensuring a single representative value for each time step.

Figure 4.2 shows the heatmap of correlation analysis of the dataset, which aims to identify and minimise multicollinearity, which is an essential step in improving the reliability and validity of predictive modelling. The study revealed significant correlations above an absolute value of 0.7 among several variables, which led to a decision to exclude features to avoid redundancy and potential bias in the model. In this analysis, it was decided to exclude *'app_temp'* (actual temperature) due to its high correlation coefficient of 0.99 with *'temp'* (temperature). This decision was made as *'temp'* measures temperature more directly. Similarly, *'solar_rad'* (solar radiation) was removed as it was highly correlated at 0.97 with *'ghi'*, which precisely measures the solar energy received by the horizontal surface and is more directly related to the analysis. The correlation between *'wind_spd'* (wind speed) and *'wind_gust_spd'* (wind gust speed) was significant at 0.8, indicating consistent measurements over time. However, *'wind_gust_spd'* was excluded to maintain *'wind_spd'* as an indicator of average wind speed.

Furthermore, it was observed that *'ghi'* showed strong correlations with *'uv'* (ultraviolet index) and *'elev_angle'* (elevation angle) at 0.94 and 0.87, respectively. *'Uv'* was removed as *'ghi'* provides a comprehensive representation of solar energy, which is central to the analysis. *'elev_angle'* was also excluded as it does not provide significant additional insight beyond what *'ghi'* already provides. These adjustments resulted in decreased multicollinearity, as shown in Figure 4.3.

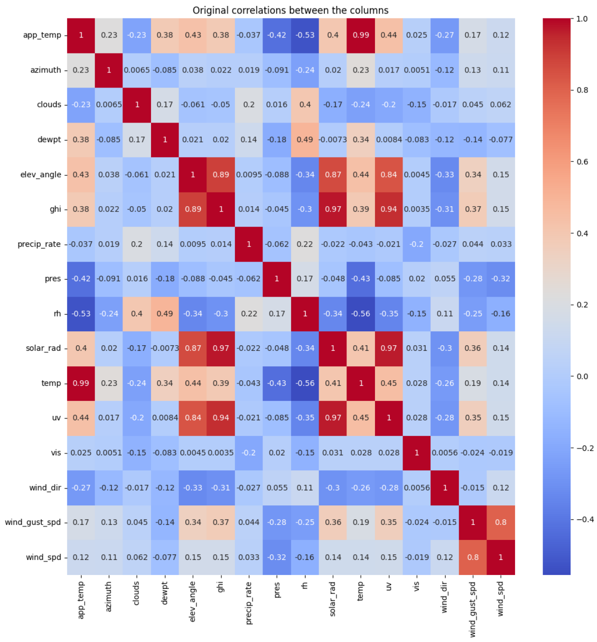

Figure 4.2: Correlation Heatmap Before Feature Selection

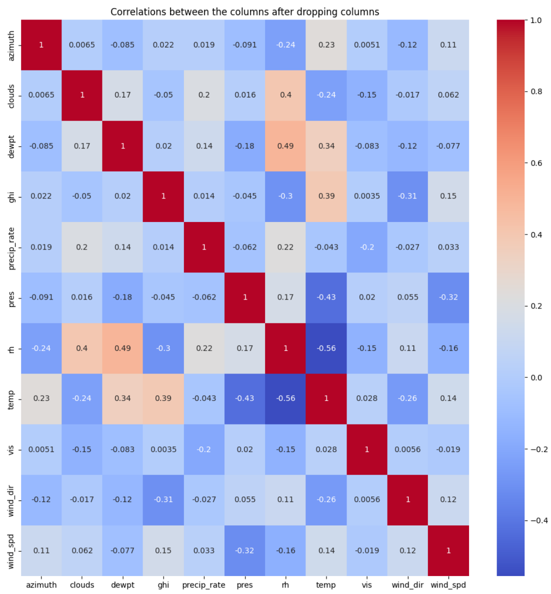

Figure 4.3: Correlation Heatmap Following Reduction

To ensure the integrity and applicability of the datasets, special attention was given to handling NaN and zero values during the refinement of the weather datasets for analysis. In particular, detecting NaN values within the *'dewpt'* (dew point) column with 94 missing items required a careful approach, filled using time-series interpolation, provided in detail in Figure 4.1.2. This methodological choice was made to maintain the temporal continuity of the data, which is crucial for weather-dependent investigations. A thorough evaluation of columns with a high incidence of zero values, such as 'cloud' (cloud coverage) and *'precip_rate'* (precipitation rate), was also conducted. This investigation, reinforced by referring to weather records source, confirmed that these zeroes were not artefacts of missing data but actual measurements indicating clear sky conditions. [53]

This project's data was segmented into training, validation, and test datasets to optimise the model's performance. The ratio also follows the 6:2:2 rule, the same method as the streetlight energy usage dataset, so the separated dataset has the same range of timestamps and data points. Table 4.3 above shows each DateTime range of the separated datasets.

Time Interpolation

The streetlight energy usage and weather datasets exhibited inconsistent documentation, although the dataset source showed a complete record. An interpolation technique was introduced using Python *Panda* library (version 2.2.0, Python version 3.12.0) to address this discrepancy. Specifically, the 'time' interpolation method was employed, which estimates missing values using the time interval between data points and applies linear interpolation based on the time axis, as stated in the official panda documentation. [44] This method preserves the unique time patterns within the dataset, providing a more accurate estimation of time series data by taking into account the time order and interval of data points.

However, it is worth acknowledging the limitations of the interpolation method. While 'time' interpolation provides the best guess based on the available data, it is impossible to fully replicate the actual unrecorded measurements, potentially leading to estimation errors in the dataset. Also, as the missing values appearing in the dataset are shown in consecutive forms, errors with the measured value may be inevitable.

4.1.3 Exploratory Data Analysis

The dataset on streetlight energy consumption was analysed to compare the daily total energy usage across various locations. Figure 4.4 shows a periodic curve pattern in the daily total usage for all streetlight locations, with the maximum usage occurring in January and the minimum in July every year. This pattern suggests that the streetlight system does not follow a uniform schedule throughout the year but rather operates differently according to the seasons. The longer daylight hours in summer mean that streetlights are on for a relatively shorter duration, whereas in winter, the shorter daylight hours result in streetlights being on longer. This implies although the streetlight system follows the fixed plan, there is possibility that is correlated with the climate change. Additionally, peaks in energy consumption are observed in some areas, indicating that the streetlights were on longer than on other days rather than being outliers. This could signify the presence of exceptional circumstances due to external factors in those areas. Lastly, the variance in the magnitude of energy consumption across different locations can be noted. This could be attributed to varying population densities, infrastructure differences within the same city, or variations in streetlight performance across different locations.

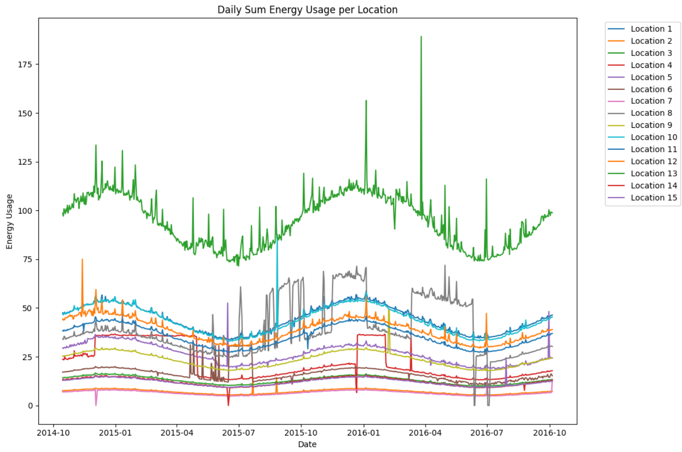

Figure 4.4: Daily total streetlight energy usage of each location

Figure 4.5 shows the time series decomposition analysis result of the variable `temp` of the weather dataset. (Refer to Appendix Figure 7.1 for the rest of the results.) This analysis reveals significant seasonal patterns for parameters such as azimuth, dew point, ghi, precipitation, and temperature. These patterns are closely linked to seasonal changes, with longer days, higher dew points in summer, colder temperatures, and increased precipitation in winter. In contrast, cloud cover and wind direction undergo unpredictable changes and are likely to be affected by random weather events such as storms. Atmospheric pressure and relative humidity trends are more consistent, showing gradual changes that reflect typical atmospheric conditions throughout the year. Visibility is an outlier that tends to decrease subtly over time, which can suggest gradual changes due to urban development or increased pollution levels. Seasonal and residual patterns in visibility data suggest that transient environmental events, such as dense fog or severe pollution incidents, can cause significant fluctuations. Therefore, it can be concluded that most weather parameters follow predictable seasonal cycles, but some exhibit complex behaviour that is shaped by environmental factors beyond seasonal changes.

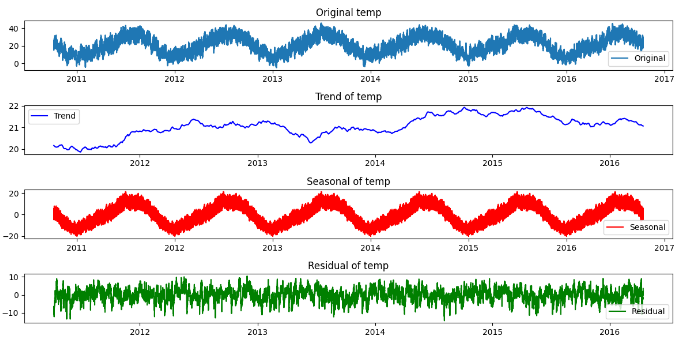

Figure 4.5: Time Series Decomposition of ‘temp‘ variable from weather dataset

Correlation analyses were conducted to verify the absence of the expected correlation between the two datasets, as mentioned in section 1.4. For this purpose, Kendall and Pearson correlation coefficients and mutual information analysis were carried out. Kendall correlation coefficient measures the ordinal association between two variables by evaluating the concordance and discordance of pairs. [27] Pearson correlation coefficient quantifies the linear relationship between two variables, with values ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation). [39] Mutual information assesses how much information one variable reveals about another, effectively capturing non-linear interactions.[25]

Figure 4.6, Figure 4.7, and Figure 4.8 illustrate the result of each correlation analysis heatmaps between the energy usage and weather datasets. The results confirmed that most weather parameters do not exhibit a significant correlation with energy usage. However, specific variables such as `wind_spd`, `pres`, and `vis` showed higher correlation coefficients than expected, which can be interpreted in several different ways. This implies that these weather conditions could potentially have a direct or indirect impact on the energy demand of streetlights. For example, strong winds or low atmospheric pressure are generally associated with unstable weather conditions, which might necessitate turning on the streetlights earlier or keeping them on for longer periods. However, as previously analyzed and acknowledged, streetlights operate strictly according to predetermined schedules, making it difficult to determine this as a meaningful correlation. Correlation does not always imply causation, and given these findings, it is challenging to confirm if the results are significant or merely coincidental. Further analysis is recommended if there is a possibility of a third-factor influence.

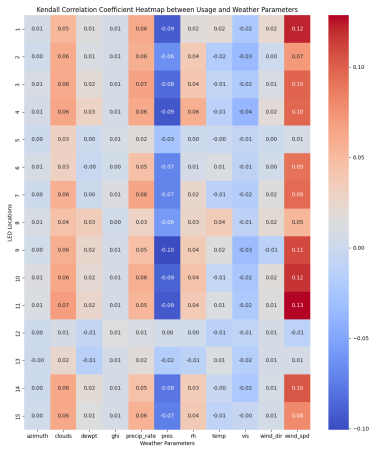

Figure 4.6: Kendall correlation coefficient heatmap

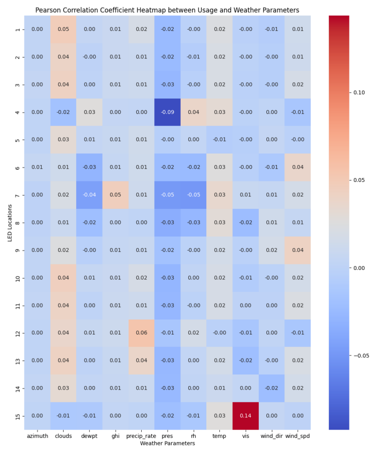

Figure 4.7: Pearson correlation coefficient heatmap

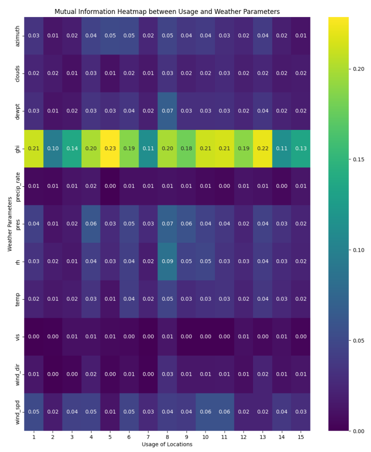

Figure 4.8: Mutual Information heatmap

4.2 LSTM Modelling

4.2.1 Modelling

Training and validation data, previously separated in the data preparation process in subsection 4.1.2, were utilised for LSTM modelling. During the data preprocessing phase, the usage column from the streetlight energy usage dataset and all parameters from the weather dataset were normalised using the *MinMaxScaler* from the *Sklearn* library (version 1.4.2). To meet the specific input and output format requirements of the LSTM model `(data size, time steps, features)` and `(data size,` `n_forecast, features)` respectively, it was necessary to transform each dataset into an appropriate form, using the timestamp as the row index. The location was used as the column index for the streetlight dataset to pivot the data so that usage values became the cell values. Similarly, the weather dataset was transformed so that each parameter became a column with the corresponding values as the cell contents. Based on this transformed data, sequence data incorporating location-specific usage and weather parameters at each timestep were generated. The LSTM was configured to produce target labels for a forecast length of 96, capturing the overall data patterns and trends over a 24-hour period and providing critical context for the DNN model. This information assists the DNN in more precisely predicting the variability over the subsequent 12 hours. Understanding long-term data trends obtained from the LSTM is useful for adjusting the input data features for the DNN, enabling it to identify and predict important patterns in narrower timeframes more accurately.

After preprocessing and transforming the datasets into a compatible format for LSTM input, the resulting data structure used for training and validation is outlined in Table 4.4 and Table 4.5 These tables show the number of samples, time steps, and features for both the streetlight energy usage and weather LSTMs. The reason why the training data points of the weather dataset exceed the other is mentioned in subsection 4.1.1. The exact values of *'timestep'* has been intentionally left unspecified in the tables, as they were subject to finalisation through this tuning process in subsection 4.2.2.

Table 4.4: Input and label structure of streetlight energy usage LSTM

Input	x_train	(41587, time_step, 15)
	x_val	(13862, time_step, 15)
Label	y_train	(41587, 96, 15)
	y_val	(13863, 96, 15)

Table 4.5: Input and label structure of weather LSTM

Input	x_train	(182573, time_step, 15)
	x_val	(13862, time_step, 15)
Label	y_train	(182573, 96, 15)
	y_val	(13863, 96, 15)

The proposed LSTM models are constructed using the *Keras* library (version 3.2.1), ensuring consistency in results through the use of random seed 42 using *Numpy* (version 1.26.4) and *Tensorflow* (version 2.16.1) library. Figure 4.9 visually represents the model's structure, showcasing every layer from input to output. The model takes sequence data through an input layer, passes it through multiple *LSTM* layers, and finally outputs predictions via a *TimeDistributed* Dense layer. Each *LSTM* layer contains a specific number of units, which is further determined in hyperparameter tuning progress (subsection 4.2.2). Besides the last layer, it is set to return sequences to pass the full data sequence to the subsequent layer. The last *LSTM* layer is set with `'return_sequences=False'`, outputting only the final time step's output to the *RepeatVector* layer. This output is then repeated `'n_forecast=96'` times by *RepeatVector* before being passed on.

Notably, this model structure differs slightly when applied to two distinct datasets. A single *LSTM* layer is used for the streetlight dataset, reflecting its relatively simple time series structure. In contrast, the weather dataset, which presents more complex patterns and temporal behaviours, employs multiple *LSTM* layers dictated by the hyperparameter `num_layers`. This adjustment in model depth according to the complexity of each dataset ensures more effective learning of the distinctive patterns inherent in each time series dataset.

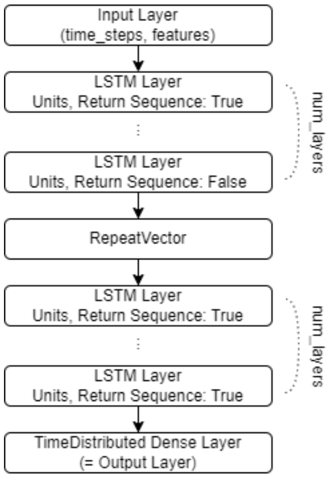

Figure 4.9: Fundamental LSTM structure of both datasets

After training the LSTM model with the prepared datasets, the model was saved for integration, and predictions were made for the validation dataset. The predictions were then rescaled back to their original scale using the inverse transformation function. The model's performance is evaluated by computing the MSE, MAE, and R^2 on these descaled predictions against the actual values. The visualised comparison with actual and predicted values and the evaluation metrics values are presented in subsection 4.2.2.

4.2.2 Hyperparameter Tuning

The most crucial part of implementing the predictive model is finding the optimal parameter combination and optimising the model's performance. The model was initially set up through hyperparameter tuning, and Bayesian Optimisation was used among various optimisation techniques.

Bayesian optimisation intelligently determines where to select the following sample (a combination of hyperparameters) using the results of previous evaluations. Table 4.6 lists all LSTM parameters relevant to the streetlight energy consumption and weather datasets. This includes parameters that require tuning as well as those that are fixed. The search space is immensely expansive given the sheer volume of potential combinations—surpassing 5,000 for streetlight usage and over 30,000 for weather data. Bayesian Optimization outperforms alternatives such as Grid Search, which thoroughly investigates every possibility, and Randomised Search, which selects combinations at random due to its broad search scope. Therefore, Bayesian Optimization was considered the most appropriate approach for this case.

Table 4.6: Hyperparameters and fixed parameters and their value ranges of street light energy usage and weather LSTM

Fixed parameters	
`features`(street light energy usage)	15
`features`(weather)	11
`n_forecast`	96
Hyperparameters	
`time_steps`	96, 100, 104, ... , 184, 188, 192
`batch_size`	16, 32, 64, 128
`learning_rate`	0.001, 0.01
`units`	50, 60, ...,100
`num_layers`(weather)	1, 2, 3, 4
`optimiser`	RMSprop, Adam, SGD, Adagrad, Adadelta, Adamax, Nadam

Figure 4.10 illustrates the fundamental flow of the Bayesian Optimization process. The process started with initial data generation, where a set of hyperparameters was selected randomly within predefined ranges above to create the first candidate model. Subsequently, the model underwent evaluation using the MSE metric. Following model evaluation, the Bayesian Optimization algorithm proceeded to update the hyperparameters. It intelligently uses the results of the previous evaluations to choose the next sample of hyperparameters, aiming to find the optimal set that minimises the MSE. The evaluation and parameter update optimisation cycle repeated iteratively, as depicted in Figure 4.10. The cycle continued until the stopping criterion of the process was reached, which, in this case, was defined by the 'max_trials' parameter of the Bayesian optimisation tuner. This parameter limits the number of trials, hence the number of hyperparameter combinations the model will experiment with to prevent time waste. Once the process identified the model with the best-performing hyperparameters, the model was saved for retraining and further integration with the DNN model.

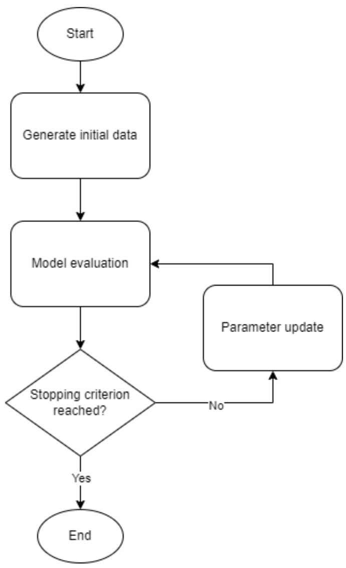

Figure 4.10: The flowchart of how the Bayesian optimisation is performed

Table 4.7 and Table 4.8 present the results of the Bayesian Optimization for the streetlight and weather datasets, detailing the hyperparameter combinations that have been identified as the most effective through the optimisation process and corresponding validation loss value, respectively.

Table 4.7: The result of the hyperparameter tuning of the streetlight dataset

Best hyperparameter combination	
`time_steps`	176
`batch_size`	128
`learning_rate`	0.001
`units`	90
`optimiser`	Adam
Validation loss	0.017414404079318047

Table 4.8: The result of the hyperparameter tuning of the weather dataset

Best hyperparameter combination	
`time_steps`	124
`batch_size`	64
`learning_rate`	0.01
`units`	50
`num_layers`	1
`optimiser`	Adam
Validation loss	0.014938274398446083

Table 4.9 compiles the performance of the LSTM models trained with these parameters in terms of MAE, MSE, and R-squared scores. Figure 4.11 and Figure 4.12 provide visualisations of the models' predictions comparison with the actual data of the first parameter with those determined parameter combinations. The comparison plots of the rest of the parameters can be found in Appendix Figure 7.2 and Figure 7.3

Table 4.9: Evaluation metrics values of the LSTM models

Evaluation Metrics	Streetlight Energy Usage LSTM	Weather LSTM
MAE	0.09088620489952391	0.07299422190210508
MSE	0.018964232276009952	0.016269708624648164
R-squared	0.6306207173611021	0.8571833935348125

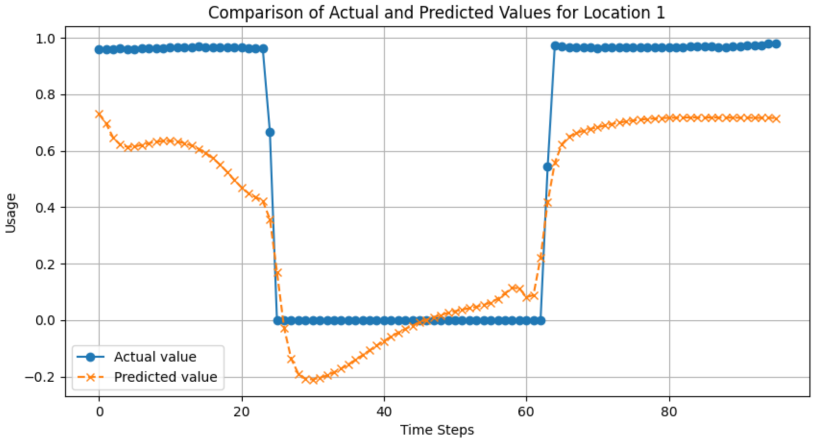

Figure 4.11: Comparison plot of the actual and predicted value of 'loc=1' from the streetlight energy usage dataset

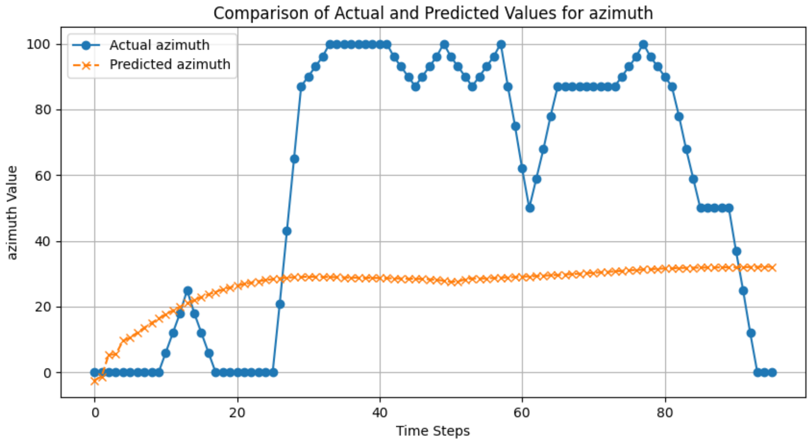

Figure 4.12: Comparison plot of the actual and predicted value of 'azimuth' from the weather dataset

4.3 LSTM-DNN Hybrid Modelling

4.3.1 Modelling

This model aims to combine the capabilities of LSTMs and DNNs to identify correlations between LED and weather data and use these correlations to predict streetlight energy usage over the next 12 hours accurately. This hybrid model structure provides two main benefits. First, it uses an LSTM to capture the temporal patterns in the time series data, and second, it uses a DNN to analyse the correlation between the two data sources to effectively understand complex time series data's hidden characteristics and patterns.

The model was developed using the *Keras* library (version 3.2.1), which used the same dataset that was used earlier to build the LSTM model. The preprocessing steps for the data were also the same as mentioned previously in section 4.2.

Figure 4.13 depicts the structure of the LSTM-DNN hybrid model. As specified in section 4.2, two pre-trained LSTM models are imported. One takes streetlight energy usage data; the other inputs weather data and outputs 96 predictions for each. These outputs are obtained from the last *LSTM* layer of the LSTM models, each of which is averaged through a *Lambda* layer. This processing is done to summarise all the information in the sequence data into a single vector, removing noise and unnecessary information from each timestep and capturing only the important features. The two reduced vectors are combined in a *Concatenate* layer, which is passed to the subsequent Dense layer. With this information, the *Dense* layer analyses the correlation between streetlight energy usage and weather. It ultimately returns a prediction of streetlight energy usage at each of the 15 locations for the next 12 hours. Through this process, the information captured by the LSTM model is simplified, and the DNN is able to perform higher-level abstractions.

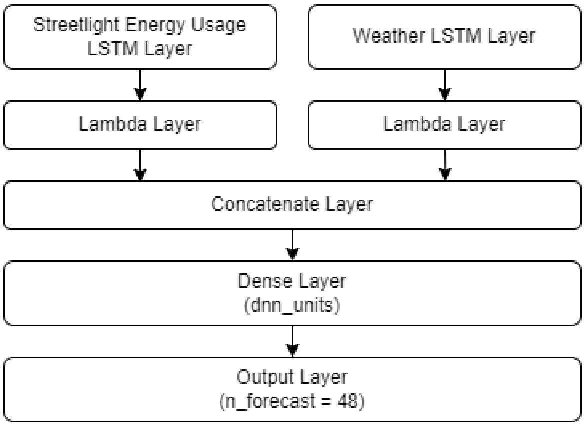

Figure 4.13: The flowchart of how the Bayesian optimisation is performed

After the LSTM-DNN model was trained using the prepared datasets, the predictions were generated for the validation dataset. These predictions were subsequently rescaled to their original scale using the inverse transformation function. The model's performance was assessed by calculating the MSE, MAE, and R^2 for these descaled predictions compared to the actual values. The visual comparisons between the actual and predicted values and the evaluation metrics are detailed in subsection 4.3.2.

4.3.2 Hyperparameter Tuning

As discussed in the independent LSTM model in subsection 4.2.2, Bayesian Optimization was chosen to navigate ample hyperparameter space shown in Table 4.10. Moreover, Bayesian Optimisation was also selected for this model given the complex interdependencies and the extensive range of potential hyperparameter combinations involved with integrating streetlight energy usage and weather LSTM.

Table 4.10: Hyperparameters and fixed parameters and their value ranges of the LSTM-DNN model

Fixed parameters	
`ts_led`	176
`ts_weather`	124
`n_forecast`	48
`batch_size`	32
Hyperparameters	
`learning_rate`	le-2, le-3, le-4
`units`	32, 64, 96, ..., 512
`optimiser`	RMSprop, Adam, SGD, Adagrad, Adadelta, Adamax, Nadam

The hyperparameter tuning of the LSTM-DNN hybrid model followed the same Bayesian Optimization process described in subsection 4.2.1 for the individual LSTM models. This method involves generating initial data with randomly selected hyperparameters, evaluating model performance using the MSE metric, and iteratively updating hyperparameters based on the outcomes. The process continues until the `'max_trials'` stopping criterion is met. For a detailed explanation of this optimisation cycle and its implementation, please refer to Figure 4.10 and the accompanying text in subsection 4.2.1.

Table 4.11 displays the outcomes of the Bayesian Optimization for both the streetlight and weather datasets used in the LSTM-DNN hybrid model. These tables detail the hyperparameter combinations found to be most effective during the optimisation process. Additionally, Table 4.12 compiles the performance metrics—MAE, MSE, and R-squared scores—of the LSTM-DNN models trained with these optimal parameters. Figure 4.14 further provides visual comparisons of the model predictions against the actual data of location 1. The comparison plots of the rest of the locations can be found in Appendix Figure 7.4.

Table 4.11: The result of the hyperparameter tuning of the hybrid model

Best hyperparameter combination	
`learning_rate`	1e-2
`units`	32
`optimiser`	Adam
Validation loss	0.021345023288055763

Table 4.12: Evaluation metrics values of the LSTM-DNN hybrid model

Evaluation Metrics	LSTM-DNN model
MAE	0.09667645333362287
MSE	0.020263945120121907
R-squared	0.6053386994994167

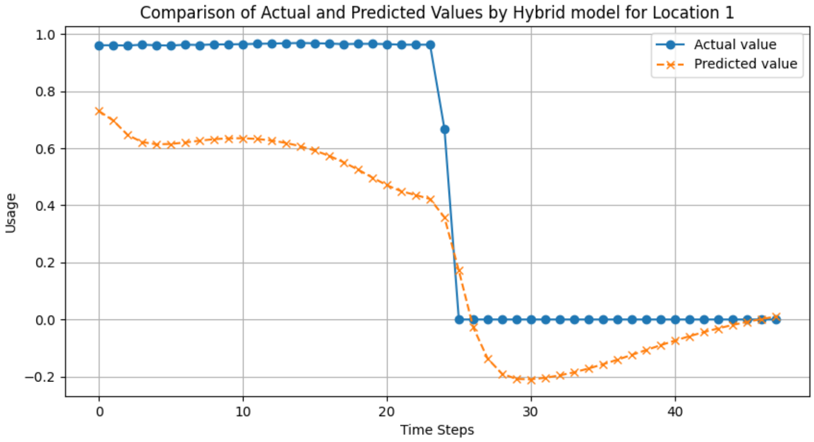

Figure 4.14: Comparison plot of the actual and predicted value by hybrid model for ‘loc=1‘

CHAPTER 5

Results & Discussion

5.1 Testing

The final evaluation using the testing data was made for the model general performance assessment. Figure 5.1 is the comparison plot of the actual and predicted values in location 1. (Refer to Figure 7.5 for the rest of the location plots) The values for MAE and MSE are 0.03277 and 0.00331, respectively as shown in Table 5.1. These figures indicate that the model has relatively low error in individual predictions. Such results can be interpreted that the model is performing precise predictions for each data point. However, the R-squared value is -0.00297, which is very low, indicating that the model almost fails to explain the overall variability of the data. An R-squared value in the negative implies that the predictions are worse than simply using the mean of the data. This means that the predictions do not effectively reflect the patterns in the actual data.

Based on these results, the LSTM-DNN model demonstrates a certain level of performance at the individual prediction level but fails to capture the overall pattern of the data. Therefore, it appears necessary to reassess the model's structure or introduce additional features for improvement. Moreover, reconsidering the data preprocessing steps to provide better input data to the model could be worthwhile.

Table 5.1: Final Evaluation metrics values of the LSTM-DNN hybrid model after testing

Evaluation Metrics	LSTM-DNN model
MAE	0.03277201051944257
MSE	0.0033061834058029332
R-squared	-0.0029742949545499275

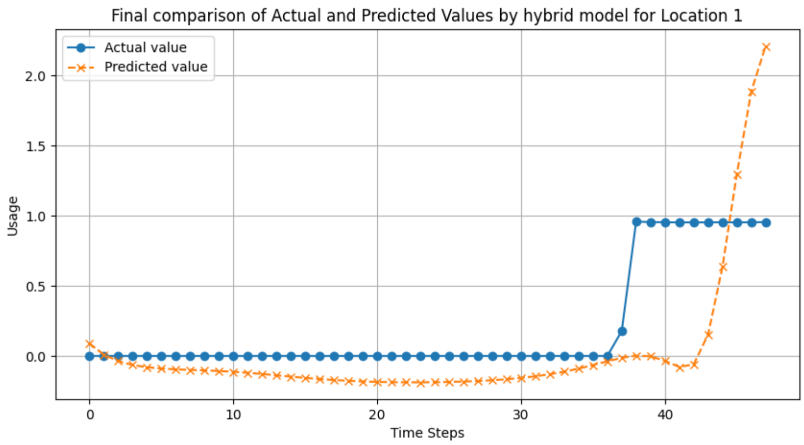

Figure 5.1: Comparison plot of the actual and predicted value by hybrid model for 'loc=1' using testing data

5.2 Limitations and Future Work

Despite the promising results obtained, several limitations exist in the current study that could impact the model's effectiveness and applicability. Firstly, hyperparameter tuning may not have sufficiently improved performance due to the limited time allocated to the process, particularly for LSTM models, which typically have longer training durations than other models. As a result, the predetermined `max_trial` value in the stopping criteria may not have been set high enough to explore the optimal configurations thoroughly. Therefore, future work should consider extending the hyperparameter tuning process without time constraints or potentially using alternative stopping criteria such as performance threshold or convergence to achieve better model performance.

Additionally, as noted in section 1.4, the streetlight energy usage dataset used in this study operates on a traditional lighting system, which turns on and off based on predetermined day and night schedules. This characteristic likely made it challenging for the LSTM-DNN model to effectively capture the correlation between energy usage and weather. To address this, future work should focus on modifying the streetlight system and improving data collection methods. For instance, integrating IoT technology to install streetlight sensors can facilitate real-time data collection on their operational status and ambient brightness. Developing a centralised data system to manage this integrated data could enable more precise analyses of energy usage patterns and their correlations with weather conditions. Such improvements could make the streetlight system more intelligent and responsive, significantly enhancing energy efficiency and contributing to the primary research goal of predicting and optimising future streetlight energy consumption.

Improving the quality and quantity of the dataset could also significantly enhance the model's performance. Including data collected from a wider and more varied geographical scope can help assess the model's generalization ability under different conditions. For example, the datasets that implies the effects of urban environmental factors or infrastructure difference can be utilised by feature engineering.

Furthermore, exploring the potential of various machine learning models is also advisable. For example, CNNs can effectively identify temporal and spatial patterns in weather data, capturing the complex interactions between weather changes and streetlight energy use. Applying CNN strengths in image processing to weather pattern recognition could allow prediction models to more accurately forecast variations in streetlight usage due to weather changes. [11] Meanwhile, GBMs, with their robust statistical learning capabilities, could model complex nonlinear relationships between multiple input variables, aiding in the detailed understanding of the connections between weather conditions and streetlight energy usage. [26] Additionally, considering an ensemble model that combines predictions from various models can be beneficial. The ensemble approach merges the strengths of individual models and compensates for their weaknesses, enhancing overall prediction accuracy. [21] This method is particularly useful when analysing the comprehensive effects of weather uncertainty and various environmental factors on energy usage. The ensemble can provide more robust and reliable prediction outcomes by learning different weather and streetlight usage data aspects. [21] Such an integrated approach is expected to improve efficiency of streetlight energy optimisation strategies significantly.

CHAPTER 6

Conclusion

The research undertaken in this project highlights the significant potential of integrating artificial intelligence, particularly LSTM-DNN hybrid models, with urban streetlight energy data to enhance energy efficiency and management in urban environments. Despite the challenges posed by the variability of streetlight usage and the predictive limitations identified through the evaluation metrics, the study has provided a powerful foundation for future advancements.

While the model demonstrated reasonable accuracy in individual predictions, its overall ability to capture the broader patterns within the data was limited, as indicated by the negative R-squared value. This outcome suggests a need for further refinement in model architecture or more sophisticated data preprocessing techniques to improve predictive performance.

Future research should focus on enhancing the data collection processes, incorporating real-time data analytics, and exploring the integration of IoT technologies for dynamic streetlight management. This approach will not only refine the accuracy of our predictions but also extend the applicability of the model to a broader range of intelligent urban planning. By advancing the methodologies used in this study, further research can significantly contribute to the development of smart city energy systems that are both sustainable and efficient.

CHAPTER 7

Appendix

7.1 Source Code

The source code can be found at: https://git.cs.bham.ac.uk/projects-2023-24/yxj183.

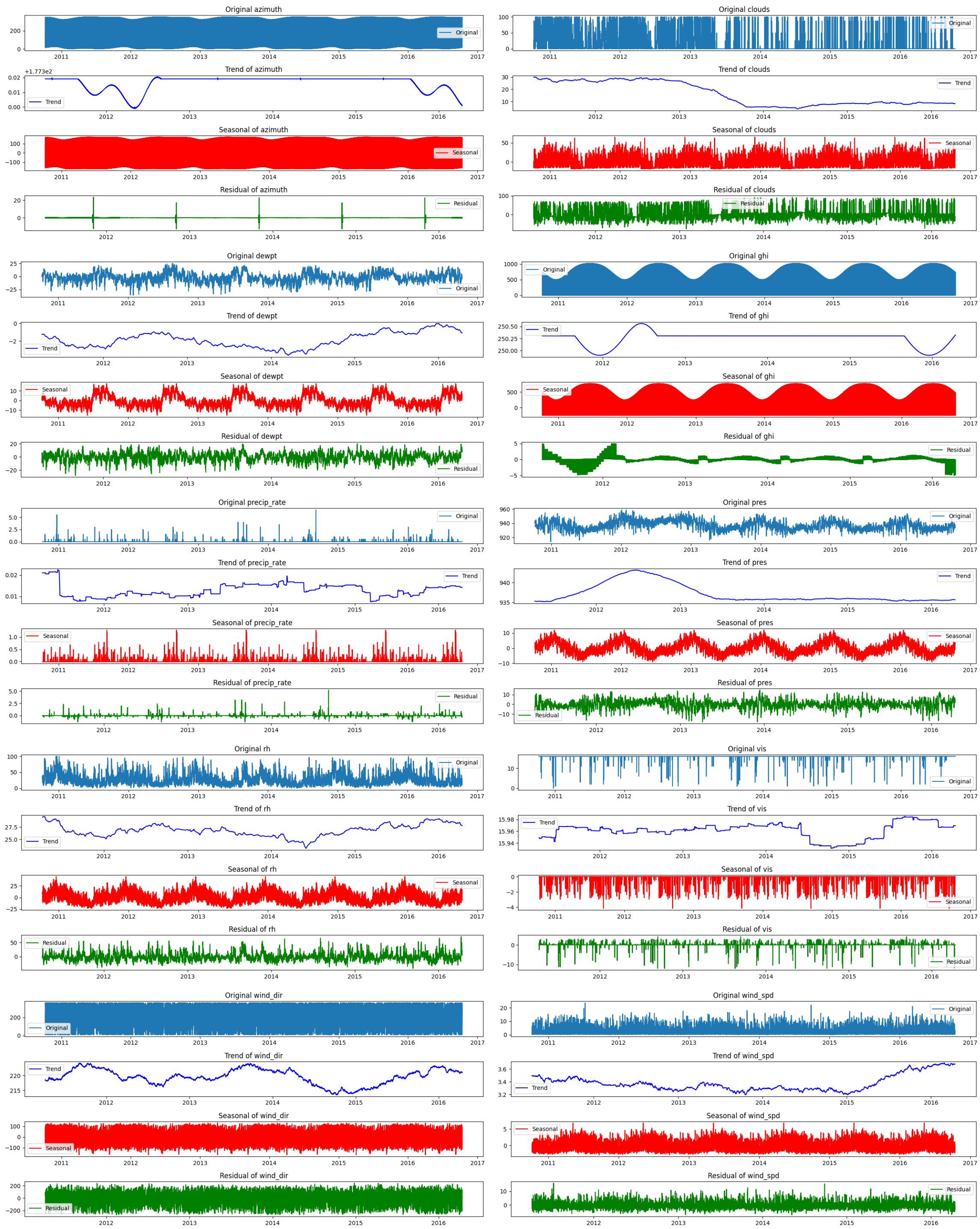

Figure 7.1: The rest of the weather attributes of the time series decomposition results

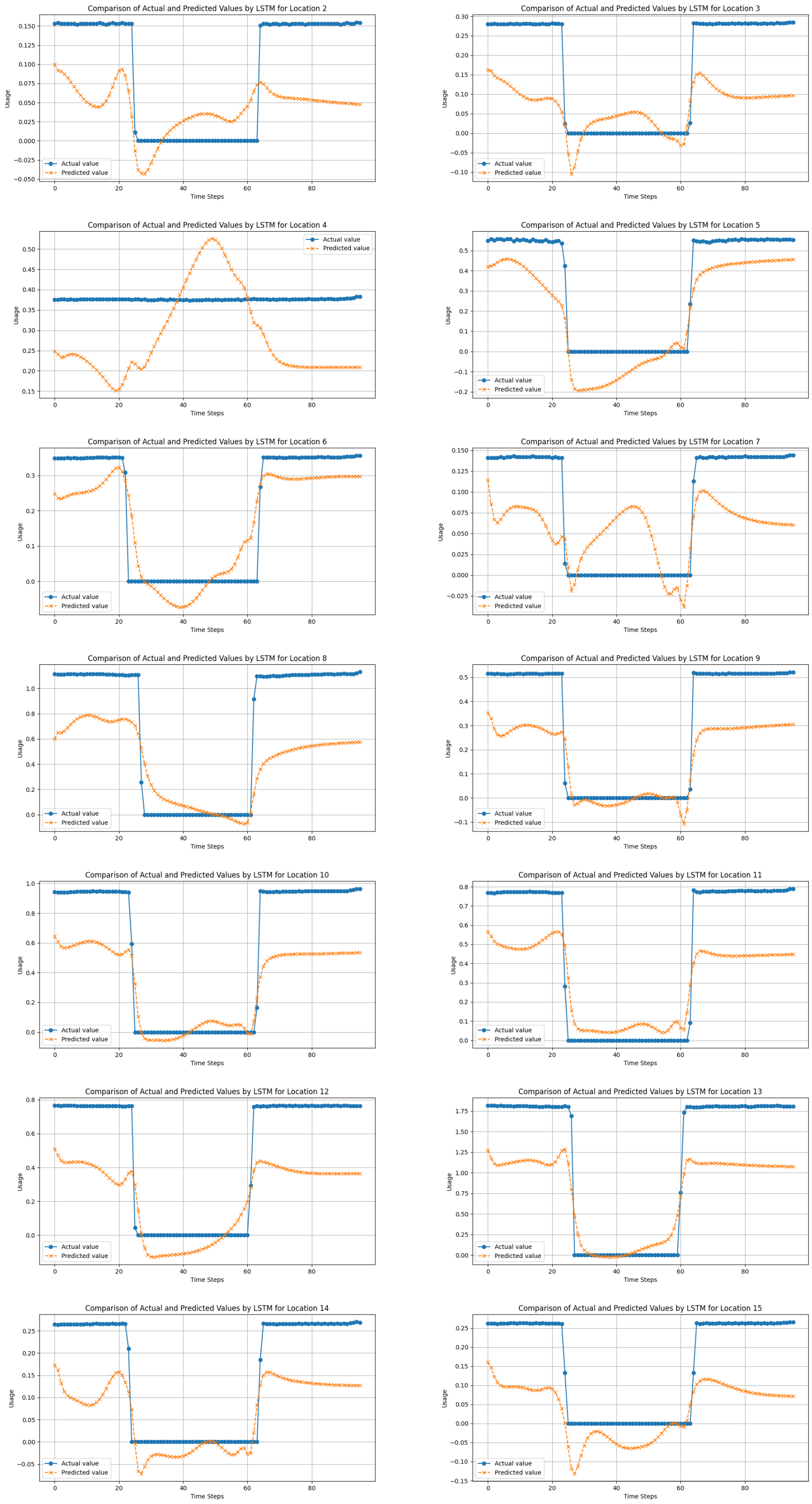

Figure 7.2: Predictions the rest of the streetlight locations by LSTM model

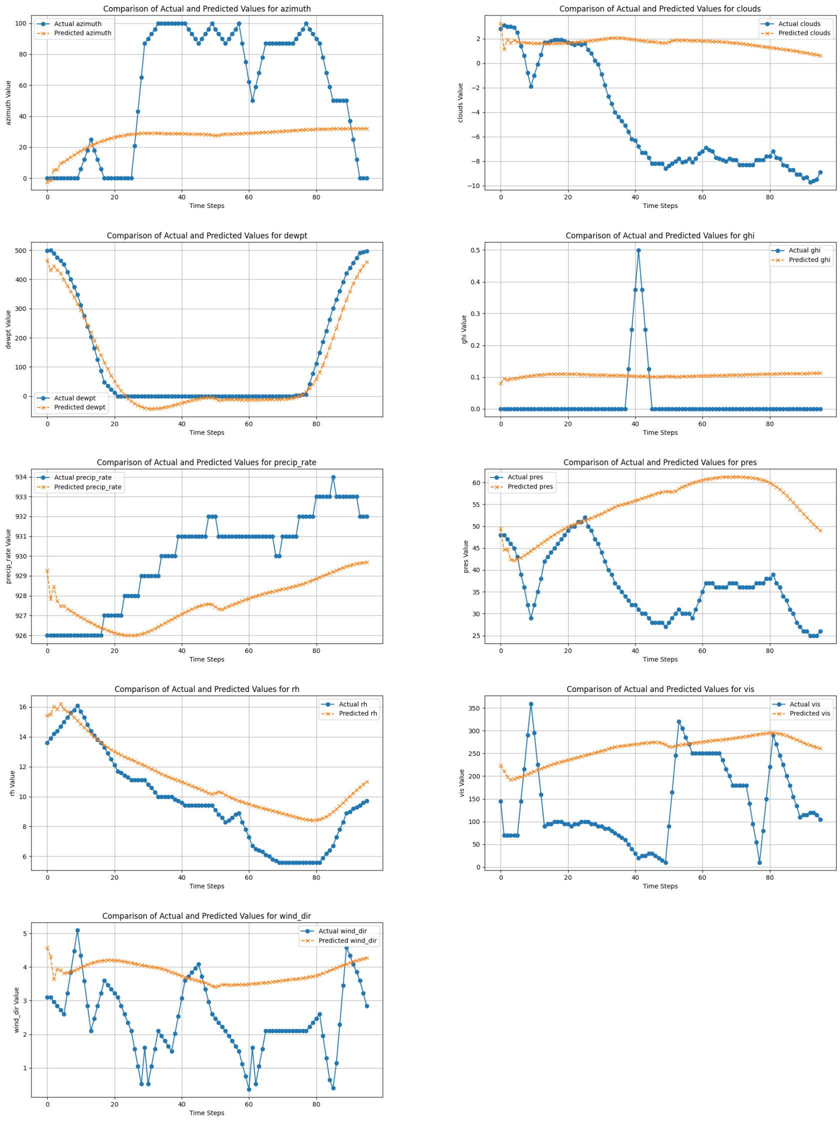

Figure 7.3: Predictions the rest of the weather parameters by LSTM model

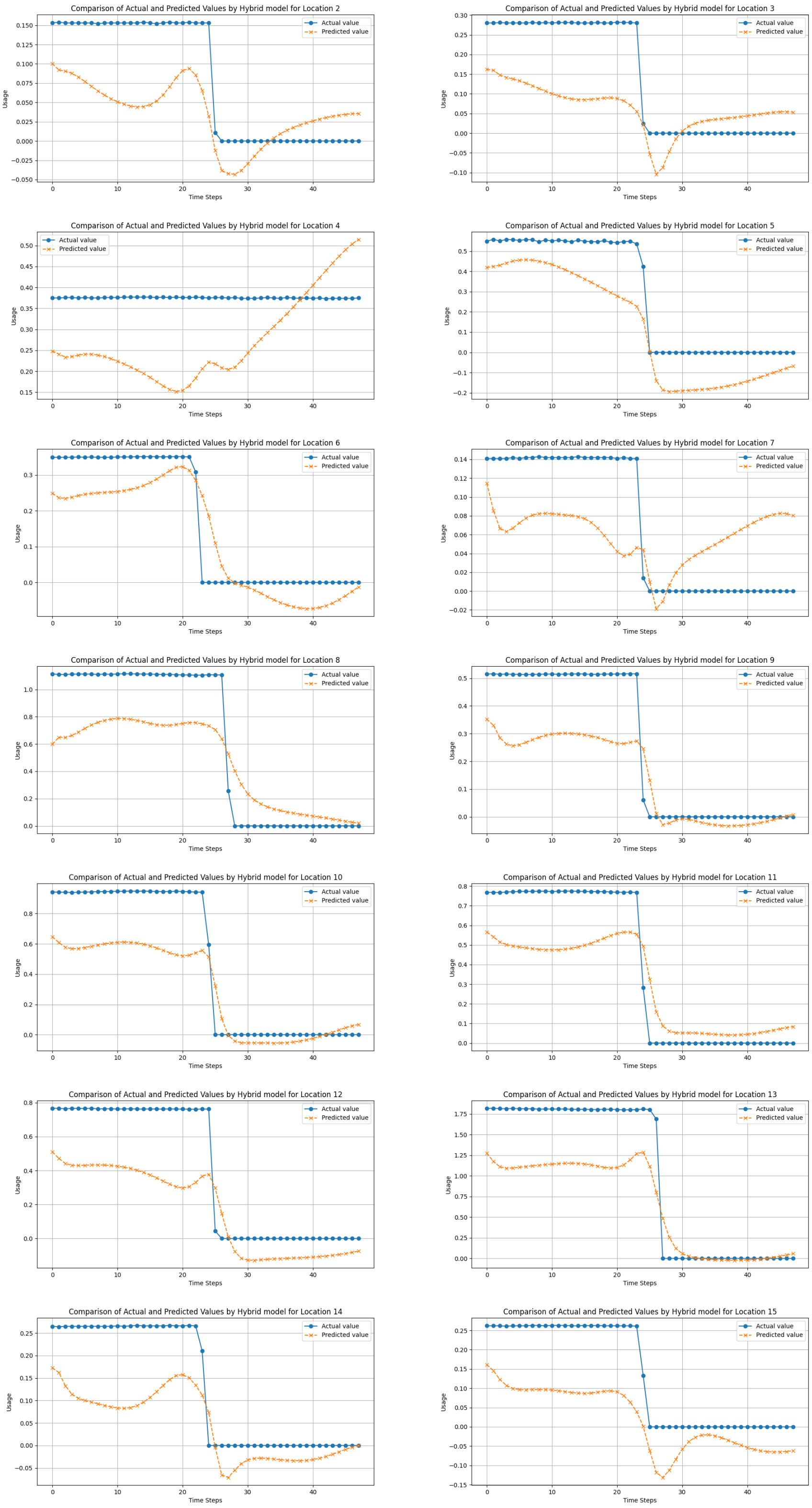

Figure 7.4: Predictions of the rest of the streetlight locations by LSTM model using the validation dataset

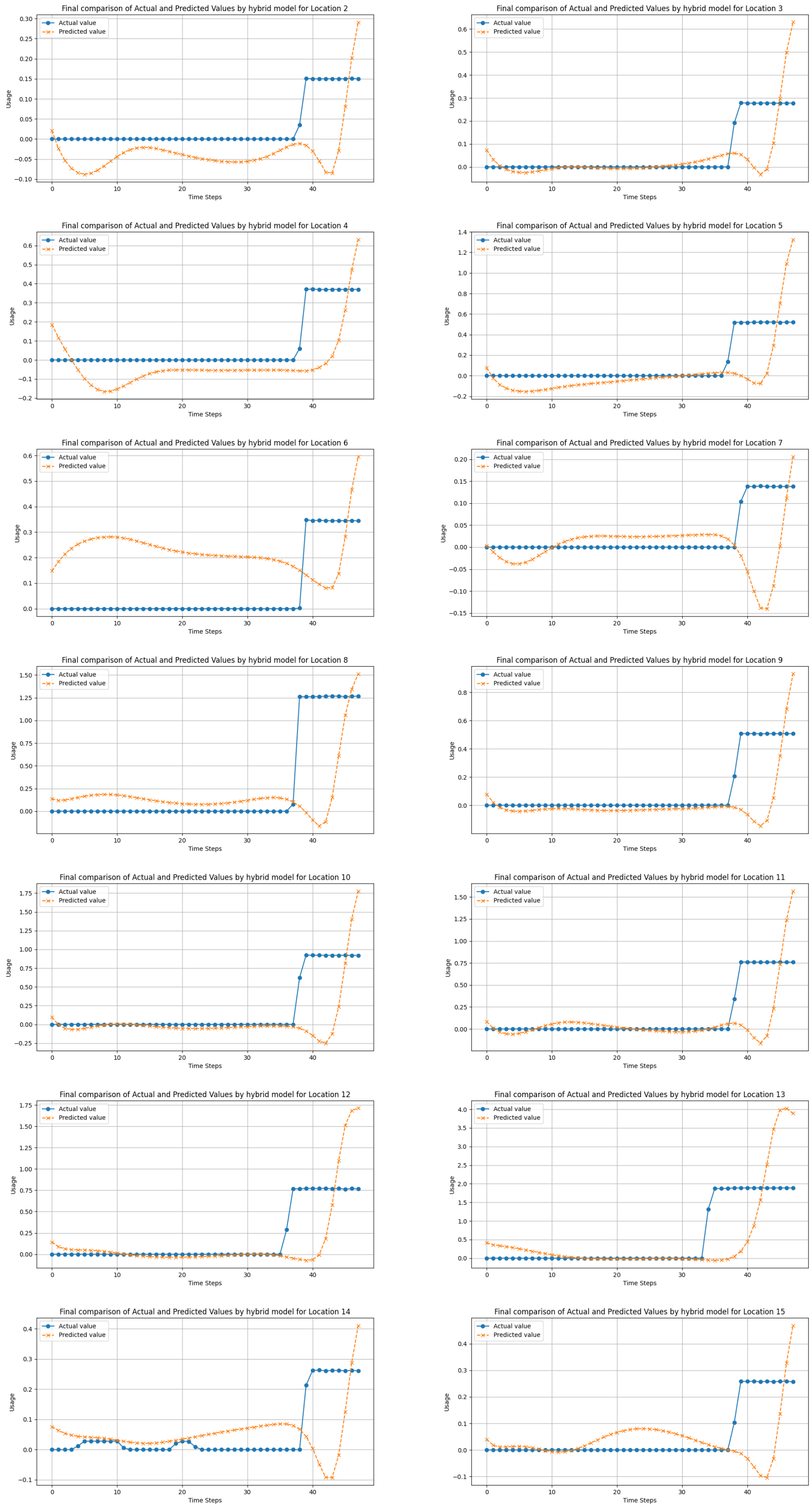

Figure 7.5: Predictions of the rest of the streetlight locations by LSTM model using the testing dataset

Bibliography

[1] M. AL-Ghamdi, A. A.-M. AL-Ghamdi, and M. Ragab. A hybrid dnn multilayered lstm model for energy consumption prediction. *Applied Sciences*, 13(20):11408, 2023.

[2] M. F. Alsharekh, S. Habib, D. A. Dewi, W. Albattah, M. Islam, and S. Albahli. Improving the efficiency of multistep short-term electricity load forecasting via r-cnn with ml-lstm. *Sensors*, 22(18):6913, 2022.

[3] K. Amasyali and N. El-Gohary. Deep learning for building energy consumption prediction. 05 2017.

[4] S. E. Bibri. A novel model for data-driven smart sustainable cities of the future: the institutional transformations required for balancing and advancing the three goals of sustainability. *Energy Informatics*, 4(1):4, 2021.

[5] C. Bikcora, L. Verheijen, and S. Weiland. Density forecasting of daily electricity demand with arma-garch, caviar, and care econometric models. *Sustainable Energy, Grids and Networks*, 13:148–156, 2018.

[6] R. E. Bird and R. L. Hulstrom. Simplified clear sky model for direct and diffuse insolation on horizontal surfaces. Technical report, Solar Energy Research Inst.(SERI), Golden, CO (United States), 1981.

[7] K. G. Boroojeni, M. H. Amini, S. Bahrami, S. Iyengar, A. I. Sarwat, and O. Karabasoglu. A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon. *Electric Power Systems Research*, 142:58–73, 2017.

[8] M. Braun, H. Altan, and S. Beck. Using regression analysis to predict the future energy consumption of a supermarket in the uk. *Applied Energy*, 130:305–313, 2014.

[9] G. Cao and L. Wu. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting. *Energy*, 115:734–745, 2016.

[10] M. Chammas, A. Makhoul, and J. Demerjian. An efficient data model for energy prediction using wireless sensors. *Computers & electrical engineering*, 76:249–257, 2019.

[11] A. Chattopadhyay, P. Hassanzadeh, and S. Pasha. A test case for application of convolutional neural networks to spatio-temporal climate data: Re-identifying clustered weather patterns. *arXiv preprint arXiv:1811.04817*, 2018.

[12] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang. Short-term electrical load forecasting using the support vector regression (svr) model to calculate the demand response baseline for office buildings. *Applied Energy*, 195:659–670, 2017.

[13] D. Chi. Research on electricity consumption forecasting model based on wavelet transform and multi-layer lstm model. *Energy Reports*, 8:220–228, 2022. 2021 International Conference on New Energy and Power Engineering.

[14] K. T. Chui, M. D. Lytras, and A. Visvizi. Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption. *Energies*, 11(11):2869, 2018.

[15] J. Chung and B. Jang. Accurate prediction of electricity consumption using a hybrid cnn-lstm model based on multivariable data. *PloS one*, 17(11):e0278071, 2022.

[16] N. City of Las Vegas. Street light energy usage. Data World, 2019.

[17] S.-L. Developers. Regression metrics — scikit-learn documentation, 2023. Accessed on: December 14th, 2023.

[18] F. Ferrini, A. Fini, J. Mori, and A. Gori. Role of vegetation as a mitigating factor in the urban context. *Sustainability*, 12(10):4247, 2020.

[19] N. Fumo and M. Rafe Biswas. Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, 47:332–343, 2015.

[20] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[21] X. Guo, X. Wang, Y. Ao, W. Dai, and Y. Gao. Short-term photovoltaic power forecasting with adaptive stochastic configuration network ensemble. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, 12, 2022.

[22] T. Han, K. Muhammad, T. Hussain, J. Lloret, and S. W. Baik. An efficient deep learning framework for intelligent energy management in iot networks. *IEEE Internet of Things Journal*, 8(5):3170–3179, 2021.

[23] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[24] W. G. Holt. Climate change impacts on urban areas. In *Urban Areas and Global Climate Change*, pages xiii–xvii. Emerald Group Publishing Limited, 2012.

[25] W. Jitkrittum, H. Hachiya, and M. Sugiyama. Feature selection via¡ mos00099. jpg¿ 1-penalized squared-loss mutual information. *IEICE TRANSACTIONS on Information and Systems*, 96(7):1513–1524, 2013.

[26] B. Kalantar, N. Ueda, V. Saeidi, K. Ahmadi, A. A. Halin, and F. Shabani. Landslide susceptibility mapping: Machine and ensemble learning based on remote sensing big data. *Remote Sensing*, 12(11):1737, 2020.

[27] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30:81–93, 1938.

[28] Z. A. Khan, T. Hussain, and S. W. Baik. Boosting energy harvesting via deep learning-based renewable power generation prediction. *Journal of King Saud University - Science*, 34(3):101815, 2022.

[29] Z. A. Khan, A. Ullah, W. Ullah, S. Rho, M. Lee, and S. W. Baik. Electrical energy prediction in residential buildings for short-term horizons using hybrid deep learning strategy. *Applied Sciences*, 10(23):8634, 2020.

[30] M. Khodayar, O. Kaynak, and M. E. Khodayar. Rough deep neural architecture for short-term wind speed forecasting. *IEEE Transactions on Industrial Informatics*, 13(6):2770–2779, 2017.

[31] T.-Y. Kim and S.-B. Cho. Predicting residential energy consumption using cnn-lstm neural networks. *Energy*, 182:72–81, 2019.

[32] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, and Y. Xu. Short-term residential load forecasting based on resident behaviour learning. *IEEE Transactions on Power Systems*, 33(1):1087–1088, 2018.

[33] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang. Short-term residential load forecasting based on lstm recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1):841–851, 2019.

[34] C. Li, Y. Tao, W. Ao, S. Yang, and Y. Bai. Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy*, 165:1220–1227, 2018.

[35] A. Mehmood, K.-T. Lee, and D.-H. Kim. Energy prediction and optimization for smart homes with weather metric-weight coefficients. *Sensors*, 23(7):3640, 2023.

[36] A. Mehmood, K.-T. Lee, and D.-H. Kim. Energy prediction and optimization for smart homes with weather metric-weight coefficients. *Sensors*, 23(7):3640, 2023.

[37] N. Mohan, K. Soman, and S. Sachin Kumar. A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. *Applied Energy*, 232:229–244, 2018.

[38] B. Nastasi and A. Mauri. Energy consumption in a smart city, 2022.

[39] R. S. B. Pearson. Book review: handbook of medical treatment. *Proceedings of the Royal Society of Medicine*, 58:217–217, 1965.

[40] M. Q. Raza, N. Mithulananthan, J. Li, and K. Y. Lee. Multivariate ensemble forecast framework for demand prediction of anomalous days. *IEEE Transactions on Sustainable Energy*, 11(1):27–36, 2020.

[41] H. Runhaar, H. Mees, A. Wardekker, J. van der Sluijs, and P. P. Driessen. Adaptation to climate change-related risks in dutch urban areas: stimuli and barriers. *Regional environmental change*, 12:777–790, 2012.

[42] V. Sangeetha and S. Andrews. *Introduction to artificial intelligence and neural networks.* MTS, 2023.

[43] H. Shi, M. Xu, and R. Li. Deep learning for household load forecasting—a novel pooling deep rnn. *IEEE Transactions on Smart Grid*, 9(5):5271–5280, 2018.

[44] P. D. Team. pandas.series.interpolate — pandas documentation. `https://pandas.pydata.org/docs/reference/api/pandas.Series.interpolate.html`, 2023. Accessed on: December 15th, 2023.

[45] G. Tsekouras, N. Hatziargyriou, and E. Dialynas. An optimized adaptive neural network for annual midterm energy forecasting. *IEEE Transactions on Power Systems*, 21(1):385–391, 2006.

[46] F. U. M. Ullah, A. Ullah, I. U. Haq, S. Rho, and S. W. Baik. Short-term prediction of residential power energy consumption via cnn and multi-layer bi-directional lstm networks. *IEEE Access*, 8:123369–123380, 2020.

[47] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca. *Python Deep Learning: Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow.* Packt Publishing Ltd, 2019.

[48] R. Vinuesa, H. Azizpour, I. Leite, M. Balaam, V. Dignum, S. Domisch, A. Felländer, S. D. Langhans, M. Tegmark, and F. Fuso Nerini. The role of artificial intelligence in achieving the sustainable development goals. *Nature communications*, 11(1):1–10, 2020.

[49] D. Vu, K. Muttaqi, and A. Agalgaonkar. A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Applied Energy*, 140:385–394, 2015.

[50] S. Wang, X. Wang, S. Wang, and D. Wang. Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting. *International Journal of Electrical Power & Energy Systems*, 109:470–479, 2019.

[51] Y. Wang, Q. Xia, and C. Kang. Secondary forecasting based on deviation analysis for short-term load forecasting. *IEEE Transactions on Power Systems*, 26(2):500–507, 2011.

[52] Weatherbit.io. Terms of service. `https://www.weatherbit.io/terms`, 2023. Accessed on: November 13th, 2023.

[53] WeatherSpark. Weather data for las vegas, nevada, united states, 2010-2016. `https://weatherspark.com/`. Accessed on: December 14th, 2023.

[54] T. Yigitcanlar and F. Cugurullo. The sustainability of artificial intelligence: An urbanistic viewpoint from the lens of smart and sustainable cities. *Sustainability*, 12(20):8548, 2020.

[55] T. Yigitcanlar, N. Kankanamge, M. Regona, A. Ruiz Maldonado, B. Rowan, A. Ryu, K. C. Desouza, J. M. Corchado, R. Mehmood, and R. Y. M. Li. Artificial intelligence technologies and related urban planning and development concepts: How are they perceived and utilized in australia? *Journal of Open Innovation: Technology, Market, and Complexity*, 6(4):187, 2020.

[56] A. Zheng and A. Casari. *Feature engineering for machine learning: principles and techniques for data scientists.* " O'Reilly Media, Inc.", 2018.

[57] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv. Vector field-based support vector regression for building energy consumption prediction. *Applied Energy*, 242:403–414, 2019.