Chapter 1

Introduction

1.1 Introduction

Forecasting is a method for predicting future occurrences by evaluating previous trends. Short-term forecasting is a method in which energy prediction is made from one hour to one week [2]. The energy consumption of a building is increased because of the use of light, air conditioning, fans, and many other electronic gadgets. As a result, buildings in such areas have increased carbon emissions compared with transportation. Hence, the prediction of the energy consumption of a building is essential for efficient energy management to reduce the carbon footprint, and it has become possible due to the concept of smart buildings [3]. A smart building is a building that can respond to the current environmental situation without human intervention. It monitors in real-time and optimizes the resources. In smart buildings, various activities are remotely controlled using Internet of Things (IoT) technology, and buildings' energy consumption is reduced significantly. Energy optimization relies on energy forecasting, which is required to avoid electricity wastage and save energy.

A smart building is defined as a building in which more than one function is performed at a time [4], and has the following features[4]:

- **Automation:** the ability to perform functions automatically.

- **Adaptability:** the ability to understand, predict, and meet user's needs.
- **Interactivity:** the ability to perform interaction among users.
- **Efficiency:** the ability to provide energy efficiency and save time and cost.

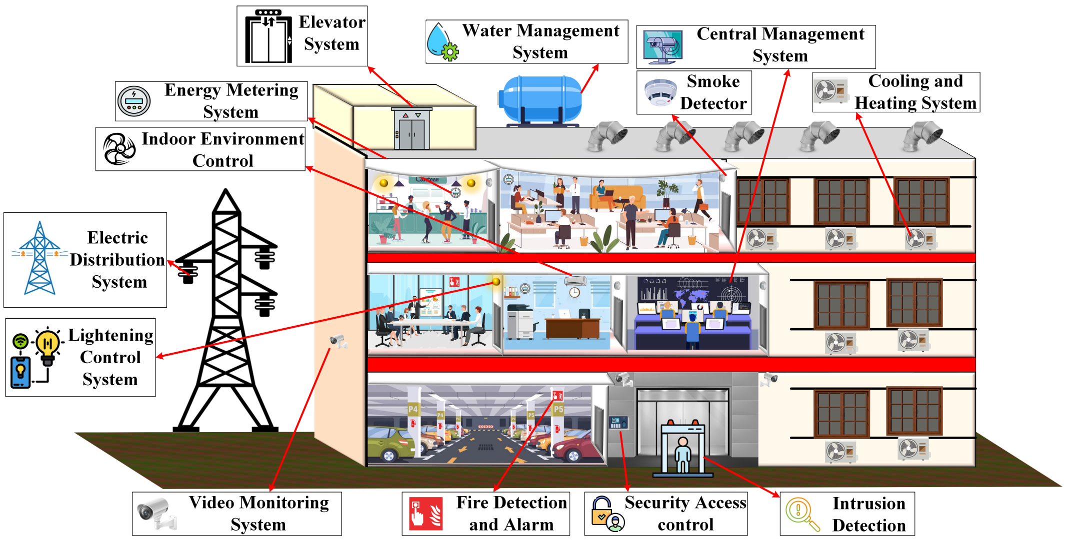

Figure 1.1: Smart Building Architecture

Smart buildings integrate advanced technologies like sensors, controls, and automation to optimize energy usage and enhance operational efficiency. Figure 1.1 shows the smart building architecture. These buildings utilize real-time data to adjust heating, cooling, lighting, and other systems based on occupancy, environmental conditions, and energy demand. Smart buildings reduce costs, improve sustainability, and enhance occupant comfort by optimizing energy consumption. They play a crucial role in energy optimization by enabling proactive energy management and demand response strategies. Smart buildings are essential for achieving energy efficiency goals, enhancing performance, and promoting environmental sustainability. Smart buildings also support the integration of renewable energy sources, contributing to a sustainable energy ecosystem. Their significance lies in minimizing energy waste, reducing carbon emissions, and increasing resilience to energy disruptions.

Short-term energy forecasting is essential for efficient energy management in various sectors, such as residential, commercial, and industrial buildings. Accurate predictions enable proactive planning and optimization of energy resources, helping to meet demand while minimizing costs and environmental impact. Organizations can make informed decisions regarding energy production, distribution, and consumption by forecasting energy consumption patterns over short time horizons, such as hours or days. So, short-term energy forecasting is essential in smart buildings for optimizing energy efficiency, reducing costs, managing peak loads, enhancing occupant comfort, integrating renewable energy, enabling predictive maintenance, and supporting integration with smart grids. Smart buildings can achieve sustainability goals by leveraging forecasted energy data and contributing to a more resilient and efficient energy system.

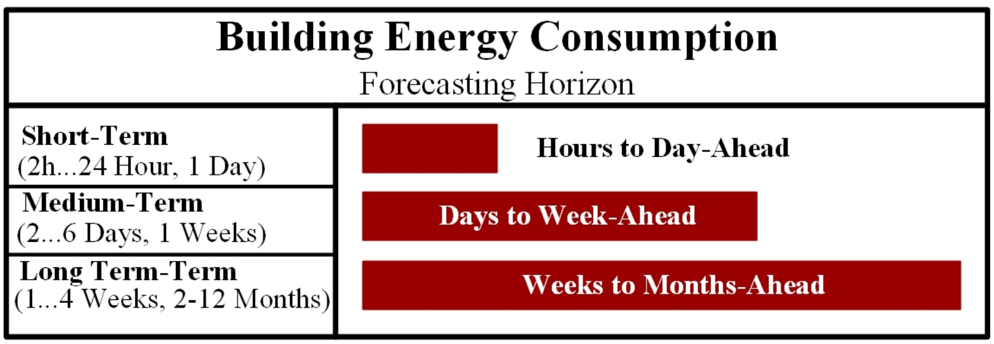

Figure 1.2: Categories of energy forecasting

The length of the future period that the forecast looks ahead to is known as the forecast horizon. The forecast horizon of energy can be categorized into short-term, mid-term, and long-term. Short-term ranges from hours to week ahead, Mid-term energy forecasting ranges from a week to a few months, and long-term forecast is made for a year [5]. The categories of forecasting horizon are explained in Figure 1.2.

1.2 Motivation

Energy demand has been increasing every day due to population growth and industrialization. This means we need more energy to meet the demand. However, traditional energy sources like coal, oil, and natural gas are bad for the Earth. They produce harmful gases that damage the environment [6]. These harmful gases pollute the environment and harm human health. Despite this, 79% of electricity still comes from fossil fuels [7]. Therefore, governments are taking steps to reduce electricity generation from fossil fuels. Now, they promote renewable or green energy [8]. It saves the environment, but production is less than the demand. Therefore, an effective energy forecasting system monitors the current energy demand of a building or industry. It supplies energy according to the building's needs and helps to reduce energy waste [9]. Similarly, it is also very useful for renewable energy management.

The waste of energy is observed due to improper use of electronic gadgets. In today's world, 40% of the world's energy consumption is covered by buildings [10]. The energy consumption of a building is increased because of the use of light, air conditioning, fans, and many other electronic gadgets. As a result, carbon emissions in buildings in such areas have increased compared to transportation. Hence, the prediction of the energy consumption of a building is essential for efficient energy management to reduce the carbon footprint, and it has become possible due to the concept of smart buildings [10]. According to a study, there were 10 million pounds of energy loss per year in the United Kingdom reported for residential buildings in 1984 due to misprediction in energy forecasting [11]. It is projected that the energy consumption of buildings will be more than 65% between 2018 and 2050 [12]. Further, Carbon Dioxide (CO_2) emissions are expected to reach 1.3 billion metric tons in 2030, which is an alarming situation for the environment and public health [12]. Therefore, accurate load forecasting is needed for smart buildings. The work done in this thesis has been focused on energy prediction in smart buildings, time series datasets, and various deep learning and machine learning approaches.

Forecasting heat energy consumption is becoming important for the effective functioning of smart buildings. It has been noted that while the demand for electricity and heat rises similarly in the winter, the demand for cooling electricity increases while the demand for heat decreases in the summer. It is difficult to store the excess heat produced because there is less demand for heat for at least three months of the year. Due to these climate features, Korea suffers from economic losses [13]. Thus, it is imperative to research heat demand and load forecasting models.

1.3 Computational aspects of data-driven techniques

This section concisely describes techniques and methodologies used to prepare the computational frameworks which are described in this thesis.

1.3.1 Machine Learning (ML) in short-term energy forecasting

Machine Learning (ML) is a subfield of Artificial Intelligence (AI), which includes the algorithms and models that allow the computer to learn from the experience without being explicitly programmed[14]. A branch of AI known as ML enables computers to automatically recognize patterns in both history and present data to generate predictions with a low-loss function. ML forecasting systems frequently offer more complex patterns and forecasting techniques. However, their main objective is to minimize the loss function and improve forecast accuracy. The sum of squares due to forecasting or prediction errors is how the loss function is commonly expressed. Most ML techniques use nonlinear approaches to minimize loss functions, whereas most conventional techniques use explainable linear processes. The ML algorithms consist of Support Vector Machine (SVM), Random Forest (RF), Linear Regression (LR), and ARIMA, which are used to forecast energy demand by considering previous data.

1.3.2 Deep Learning (DL) in short-term energy forecasting

Deep Learning (DL) is an ML subfield based on the Artificial Neural Network (ANN) [15]. An AI-based model called ANN was motivated by biological neural networks. ANN models have generally been applied to analyze nonlinear and complicated issues.

ANN models provide appropriate answers because of their high degree of self-learning, flexibility, and nonlinearity. ANN looks for patterns in datasets. An ANN consists of interconnected neurons, input nodes, hidden units, output units, iterations, number of clusters, and learning models. Two key components of deep learning data processing are data normalization and cleaning. Completing missing values and removing inappropriate values from the data set are two steps in the data-cleaning process. Data preparation can accelerate the gradient descent convergence rate, which improves prediction performance models. Gradient descent can become very complicated when dealing with non-standard data. The performance of the suggested model can be improved by normalization, which can assist deep learning models in extracting improved features from past data. DL models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), etc., can learn complex patterns in the dataset, which helps to extract the data's important features to forecast future trends in terms of the hour of the day, day of the week, etc. The LSTM model has become one of the most widely used in DL in recent years. The LSTM layers preserve the forward and backward propagated weights. LSTM uses gate control to combine long- and short-term memory. Like neural networks, deep learning models can capture non-linear correlations between the input data and energy production or consumption.

1.3.3 AutoRegressive Integrated Moving Average (ARIMA)

The AutoRegressive Integrated Moving Average (ARIMA) model combines autoregression and moving average methods. The 'AR' part predicts future values based on past data [16]. 'MA' calculates errors from past errors. The 'I' part involves making data stationary through differencing. This is necessary for ARIMA modeling. ARIMA is useful for analyzing past data and predicting future points. It's a helpful tool for forecasting time series data. ARIMA models are widely used in various fields. They are effective for fitting historical data and making predictions. By understanding past patterns, ARIMA can forecast future trends. It is an essential tool for decision-making and planning. A nonseasonal ARIMA model is known as an "ARIMA(p,d,q)" model, where: p represents the number of autoregressive terms, d signifies the number of nonseasonal differences required to achieve stationarity and q indicates the number of lagged forecast errors in the prediction equation Mathematically, ARIMA (p, d, q) can be represented as follows:

$$ARIMA(p,q) = \gamma_0 + \gamma_1 z_{t-1} + \gamma_2 z_{t-2} + ... + \gamma_p z_{t-p} + \eta_t - \delta_1 \eta_{t-1} - \delta_2 \eta_{t-2} - ..... - \delta_q \eta_{t-q} \tag{1.1}$$

Where z_t is dependent variable at time t, z_{t-1}, z_{t-2}, ... z_{t-p} are dependent variable at different time lags, γ_0, γ_1 , γ_2 ,.....γ_p and δ_1, δ_2 ,....δ_q are estimated coefficient of AR(p) and MV(q) respectively, η_t is error term at time t, η_{t-1}, η_{t-2},.....η_{t-q} are error term of previous time periods.

1.3.4 Linear Regression (LR)

Linear Regression (LR) analysis helps to predict one variable's value based on another variable [17]. The variable being predicted is the dependent one. The variable used for prediction is the independent one. This analysis finds coefficients for the linear equation, which includes independent variables to predict the dependent one best. It fits a line or surface that minimizes differences between expected and actual values. Simple linear regression calculators find the best-fit line using the "least squares" method. Then, estimate the dependent variable's value from the independent one. Linear regression is widely used in various fields for prediction and analysis tasks. The mathematical way of representing the linear regression is as follows:

$$p = \beta_0 + \beta_1 q \tag{1.2}$$

p is the dependent variable q is the independent variable β_0 is a constant. β_1 is the regression coefficient, and it is defined as follows:

$$\beta_1 = \frac{\sum[\frac{q_i - q}{p_i - p}]}{\sum[q_i - q]^2} \tag{1.3}$$

where q and p are mean value and q_i and p_i are observed data points

1.3.5 Long Short-Term Memory (LSTM)

LSTM is a recurrent neural network that can understand the order of events in sequence prediction tasks [18]. They are good at solving complex problems like machine translation and energy consumption prediction. Unlike RNNs, which struggle with long sequences, LSTM can handle long-term data. Additionally, LSTM addresses the vanishing gradient problem encountered by RNNs during training on long sequences. It remembers previous long-term data and decides which information to keep or discard using gates. These gates include the input gate, forget gate, and output gate. The input gate controls new information entering the memory. The forget gate removes irrelevant information from memory. The output gate decides what information to use for the next step. LSTM can automatically manage stored memory using these gates, making it effective for long-term data analysis. Mathematically, input gate, forget gate, output gate, cell state, candidate cell state, and final output state are represented as follows:

$$i_t = \sigma_g(W_i[h_{t-1}, x_t] + b_i) \tag{1.4}$$

$$f_t = \sigma_g(W_f[h_{t-1}, x_t] + b_f) \tag{1.5}$$

$$O_t = \sigma_g(W_O[h_{t-1}, x_t] + b_O) \tag{1.6}$$

$$\tilde{c}_t = \sigma_c(W_c[h_{t-1}, x_t] + b_c) \tag{1.7}$$

$$c_t = f_t * c_{t-1} + i_t + \tilde{c}_t \tag{1.8}$$

$$h_t = O_t * \sigma_c c_t \tag{1.9}$$

where σ, w_x, h_{t-1}, x_t, b_x are sigmoid functions, weights of different gates of neurons, previous output, input, and bias of different gates respectively.

1.3.6 Gated Recurring Units (GRU)

The Gated Recurrent Unit (GRU)[19] is a type of RNN that speeds up LSTM networks when handling large amounts of data. GRU simplifies the LSTM structure to make processing faster. LSTM and GRU have strengths, but if speed is crucial, GRU is preferred, while LSTM is better for accuracy. Choosing between them depends on what is most important for the task. GRU has two gates: the reset gate and the update gate. The reset gate helps adjust how new input combines with memory, while the update gate controls how much old memory is kept. These gates allow each hidden unit to decide what to remember or forget as it processes a sequence. Mathematically, the GRU can be expressed as follows: The update gate (z) and reset gate (r)are calculated using both the current input (x) and the previous hidden state (h_{t-1}).

$$Z_t = \sigma(W^z.[h_{t-1}, x_t]) \tag{1.10}$$

$$R_t = \sigma(W^R.[h_{t-1}, x_t]) \tag{1.11}$$

The candidate activation vector ($\tilde{h_t}$) and hidden state (h_t) are calculated using the current input (x) and a version of the previous hidden state that the reset gate adjusts and determined by merging the candidate activation vector with the prior hidden state, adjusted by the update gate's weight respectively as follows.

$$\tilde{h_t} = tanh(w^h.[R_t * h_{t-1}, x_t]) \tag{1.12}$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * h_t \tag{1.13}$$

where the matrices W^z,W^R and w^h are weights matrices the model learns during training.

1.3.7 Bidirectional Encoding Representations from Transformers (BERT)

BERT is built on top of the Transformer architecture, which is well-known for easily handling interdependencies in sequential data. With the help of multiple layers of self-attention processes, the model determines the relative value of different words in a sentence. The masked language model (MLM) and next sentence prediction (NSP) are two unsupervised tasks used to pre-train it on a large amount of textual input. Unlike previous language models, BERT processes sequences considering the previous and subsequent contexts. It may successfully extract deeper meanings and capture a more comprehensive context because of its bidirectional nature. BERT can be optimized for the final tasks, such as sequence classification and energy consumption prediction of time series data.

1.4 Objectives

The objectives of this thesis are as follows:

- To propose a short-term energy forecasting approach for smart buildings.
- To propose an energy prediction approach with occupant count for an Office building by considering a three-year dataset of building energy management and occupancy data for smart buildings.
- To design a hybrid deep learning-based VMD-LSTM model to forecast energy based on occupancy count for the American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE) global building occupant behavior dataset for smart building.
- To develop an energy prediction approach in extreme weather conditions for smart buildings.

- To design and develop an intelligent heat energy consumption prediction model for smart buildings.

1.5 Contributions

This thesis has five contributing chapters. A brief description of each chapter is as follows:

- **Chapter-2**: Literature Review

  This chapter performs the literature review and discusses the state-of-the-art techniques and research gaps.

- **Chapter-3**: Short-Term Energy Forecasting for Smart Buildings.

  This chapter proposes a novel hybrid deep learning-based model for energy forecasting techniques for smart buildings. The proposed model is analyzed for short-term forecasting using two variants of RNN, i.e., LSTM and GRU. The proposed model has been evaluated and compared on a publicly available dataset with other alternative approaches on error metrics like MAE, RMSE, and R^2 Score. The experimental results show that the proposed technique is more superior as compared to existing state-of-the-art techniques for energy forecasting in smart buildings.

- **Chapter-4**: Energy Prediction Mechanism with Occupant Count Using CNN-BiLSTM Hybrid Model for Smart Buildings

  This chapter presents a hybrid architecture of CNN and BiLSTM for predicting energy consumption. It jointly considers both the occupancy count and the timestamp information. The proposed model considers spatial and temporal dependencies that significantly increase prediction accuracy. The performance of the proposed model is assessed using real-world data. Furthermore, the accuracy and effectiveness of the model are demonstrated in comparison to existing state-of-the-art methods in terms of predicted accuracy, MAE, and MSE. The experiment uses a novel dataset of building energy consumption and occupant count to test our hybrid model. This work elucidates energy usage patterns and enhances decision-making for energy management.

- **Chapter-5**: Occupancy-based Short-Term Energy Prediction using VMD-LSTM Hybrid Model

  This chapter details a hybrid deep learning-based VMD-LSTM model that has been proposed for forecasting the energy of a building based on the occupancy count. The model is analyzed for energy prediction based on occupancy count in the different lag environments, such as minute-wise, 24-hour, or day-wise for smart buildings, and forecasting energy for different time intervals in the coming days. The proposed model has been evaluated and compared with other alternative approaches on error metrics like MAE, RMSE, MAPE, and R^2 score.

- **Chapter-6**: Energy Prediction using Extreme Weather Conditions.

  This work is about maximizing energy consumption and guaranteeing sustainable energy storage; energy forecasting is crucial in smart buildings. This research presents a novel method for energy consumption prediction in smart buildings using the Bidirectional Encoder Representations from Transformers (BERT) model. The study investigates the relationship between weather and energy use by examining six years' worth of data from the 'Manufacture of Leather and Related Products' industry. Achieving better outcomes than all other models, the BERT model using multivariate data has an MAE of 0.011, an MSE of 0.002, a MAPE of 0.070, and a R^2 of 0.979. These results highlight how meteorological data can be used to estimate energy use more precisely. This is a useful tool for smart building energy optimization.

- **Chapter-7**: Intelligent Modeling and Prediction of Heat Energy Consumption in Smart Buildings

  This chapter details how energy demand forecasting is essential for the efficient operation of smart buildings. It has been observed that the demands for heat and electricity increase to a similar degree during winter. On the other hand, electricity demand for cooling surges, whereas heat demand plummets in summer. The significant contribution of this chapter is that an efficient hybrid deep learning model has been developed by utilizing the capabilities of CNN and RNN to predict heat energy in smart buildings. These two models are merged in parallel to capture spatial and temporal relationships of data points to forecast heat energy. Smart heat meters were installed in Danish residential buildings for data collection for three years. In this work, two scenarios are used for analysis: 1) Short-term heat energy demand prediction for particular types of houses and 2) Short-term heat energy demand prediction for individual houses. Using a Danish residential building dataset, we comprehensively evaluate the proposed model over the prediction of 1-day and 7-day ahead time horizons. The results show that the proposed model outperforms as compared to the other state-of-the-art methods in terms of RMSE, MAE, R2, and MAPE. The overall layout of the thesis has been shown in Figure 1.3.

1.6 Thesis Organization

The chapters of this thesis are summarized as follows:

Chapter 1 presents recurrent neural network (RNN) and convolutional neural network (CNN) capabilities to create an effective hybrid deep learning model that forecasts the hourly energy consumption in smart buildings. Two RNN variants such as, the Gated Recurrent Unit (GRU) and the Long Short-Term Memory (LSTM), are used to analyze the model for short-term forecasting. Experimental results show that the proposed CNN-Gated Recurrent Unit (GRU) model performs better than the other state-of-the-art methods.

Chapter 2 consists of the CNN and BiLSTM hybrid architecture for energy con-

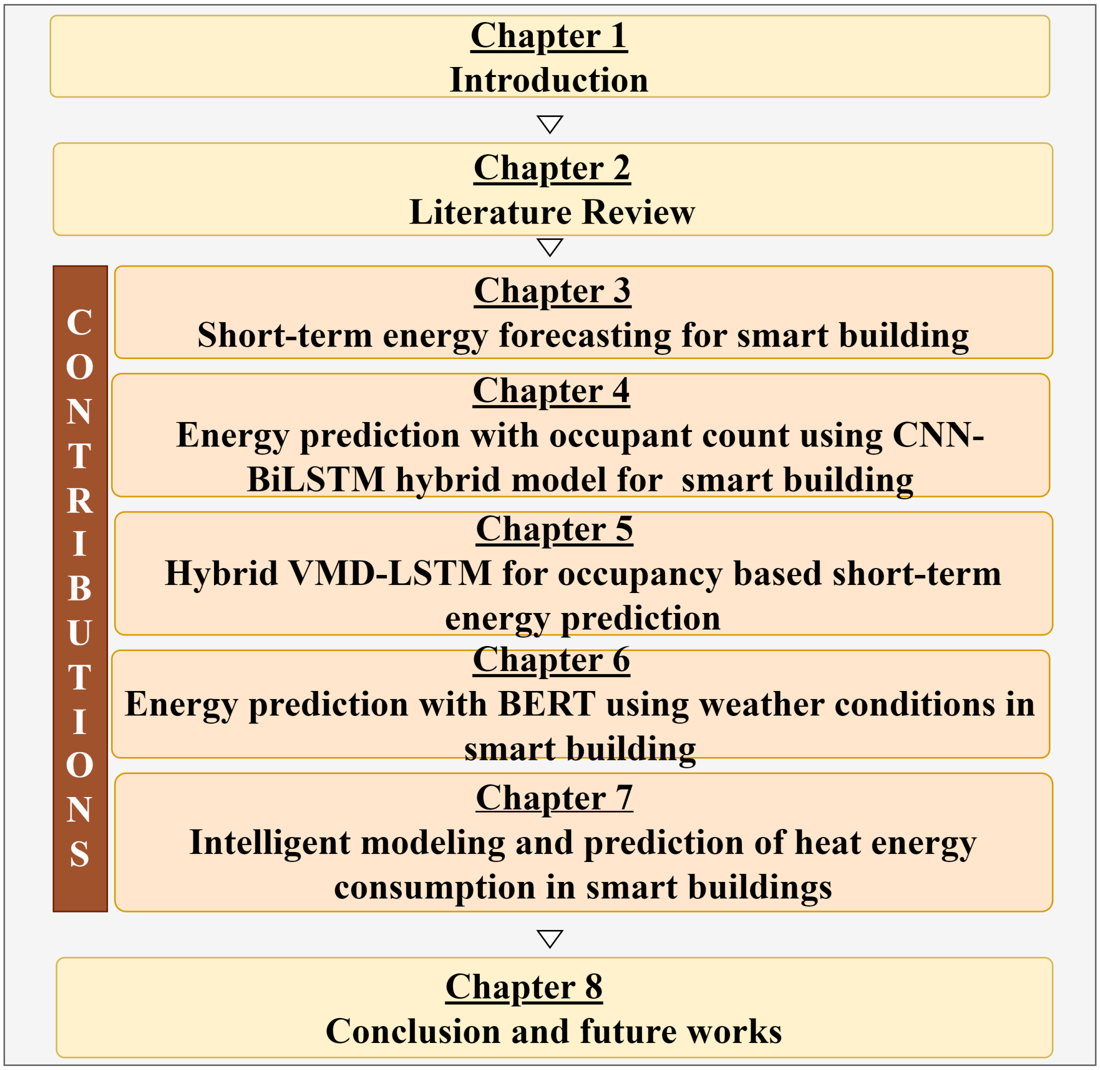

Figure 1.3: Layout of the Thesis

sumption prediction that considers the timestamp information and the occupancy count. Prediction accuracy is much improved by the suggested model, which considers temporal and geographical correlations. Predicted accuracy, MAE, and MSE illustrate the model's effectiveness and accuracy about current state-of-the-art techniques. This chapter clarifies trends in energy consumption and improves energy management decision-making.

Chapter 3 This chapter proposes a hybrid deep learning-based Variational Mode Decomposition (VMD)-LSTM model for building energy forecasting based on occupancy count. The model is examined for projecting occupancy counts in various lag contexts, including minute-by-minute, 24-hour, or day-by-day for smart buildings, and energy forecasting for various lags in the upcoming days. The proposed model has been compared with other state-of-the-art methods using error metrics such as R^2 score, RMSE, MAE, and MAPE.

Chapter 4 describes the Bidirectional Encoder Representations from Transformers (BERT) model as an innovative approach for predicting energy consumption in smart buildings. Using data from the 'Manufacture of Leather and Related Products' industry spanning six years, the proposed scheme looks into the connection between weather and energy consumption. This chapter compares the predictive power of four machine learning models using MAE, MSE, MAPE, and R2 metrics for 1, 8, and 24-hour energy consumption intervals.

Chapter 5 considers the Danish residential buildings dataset for three years. This chapter shows two scenarios 1) Short-term heat energy demand forecast for specific house types and 2) Short-term heat energy demand forecast for individual residences. To evaluate the performance of the proposed scheme, the following features, such as the minimum, maximum, and average consumption within a certain duration and the current hour, weekday, and month, are considered. We thoroughly evaluate the proposed model over the prediction of 1-day and 7-day forward time horizons using a dataset of residential buildings in Denmark. The outcomes demonstrate the effectiveness and good performance of the proposed model, which beats other cutting-edge techniques in terms of RMSE, MAE, R2, and MAPE.