**CIV E 779 Project Report**

Alexander Jordan (ID: 1577190)

Bowen Yang (ID: 1606931)

**Real-Time and Future Building Energy Prediction Using Machine Learning Techniques**

# Introduction

## Background

Building sectors account for 30%-40% of global energy consumption, contributing approximately 19% of greenhouse gas emissions [1]. Heating, ventilation, and air conditioning (HVAC), domestic hot water, lighting, and appliances are the primary services that consume energy in buildings. Globally, HVAC systems account for approximately 40% of total energy consumption [2], and a growing population and rapid urbanization are responsible for the increase in energy consumption. One of the essential parts of building energy management is energy consumption prediction, which could inspire energy policy to reduce energy consumption [3]. Furthermore, daily energy management relies on the energy demand forecast to control the appropriate energy-related equipment. For example, predicted model can be fed to a control system to pre-heat or pre-cool a building before tenants arrives or to storage generated electricity from solar panels if it predicts a peak in electrical usage. Thus, the energy prediction is a key to enable into smart and sustainable buildings design for both new and existing buildings. Much efforts have been made to accurately predict building energy consumption and design an optimal energy efficiency control system to reduce greenhouse gas emissions by managing energy emissions and conserving energy in buildings.

## Problem description

Numerous variables, including weather, time, building occupancy, etc., make it difficult to develop a reliable energy prediction model. Machine learning (ML) models, including tree algorithms, support vector machine (SVM), random forest (RF), artificial neural networks (ANN) are able to describe the relationship between model input and output without the need for complex domain knowledge, have become more and more important in building energy prediction system. But the traditional ML techniques that mentioned above have limited capacity to deal with the time-series data in predicting future energy consumption (i.e., next hour, next day and next month). Even though recurrent neural networks (RNN) performs well in tackling the sequentially data, it has the limitation in capturing long-term dependencies in the sequential data [4]. Long short-term memory (LSTM) models, as a variant of the RNN, can learn the long-term dependence information to achieve good prediction results.

## Data description and visualization

The main goal of this project is to estimate the real-time and future total energy consumption of a smart building using machine learning approaches.

The dataset (https://www.kaggle.com/code/malekzadeh/smart-home-data-processing-weather-vs-energy/data) is fetched from Kaggle. The data collection period was from 2016-01-01 to 2016-12-16 of a smart house. After removing some duplicated and irrelevant features based on authors’ experience, nineteen features including solar generation, appliance (dishwasher, furnace, home office, fridge, garage door, kitchen, barn, microwave, and living room), and weather data (temperature, humidity, visibility, pressure, windspeed, and dew point with 1min time interval are extracted to predict real-time and future (next hour) total energy consumption. The information of measured variables is shown in Table 1.

**Table 1** Information of variables.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Categories | Variables | Type | The minimum | The maximum | Unit |
| Energy generation | Generation | Numerical | 0 | 613.88 | W |
| Energy consumption | Dishwasher | Numerical | 0 | 1401.77 | W |
| Furnace | Numerical | 0.33 | 2472.63 | W |
| Home office | Numerical | 0.08 | 971.75 | W |
| Fridge | Numerical | 0.06 | 851.27 | W |
| Wine cellar | Numerical | 0.02 | 1273.94 | W |
| Garage door | Numerical | 0.02 | 1088.93 | W |
| Kitchen | Numerical | 0 | 2265.87 | W |
| Barn | Numerical | 0 | 7027.90 | W |
| Well | Numerical | 0 | 1633.02 | W |
| Microwave | Numerical | 0 | 1929.80 | W |
| Living room | Numerical | 0 | 465.22 | W |
| Weather condition | Temperature | Numerical | -12.64 | 93.72 | °F |
| Humidity | Numerical | 0.13 | 0.98 | % |
| Visibility | Numerical | 0.27 | 10 | Kilometers |
| Apparent temperature | Numerical | -32.08 | 101.12 | °F |
| Pressure | Numerical | 986.4 | 1042.46 | Millibar |
| Windspeed | Numerical | 0 | 22.91 | Km/h |
| Dewpoint | Numerical | -27.24 | 75.49 | °F |
| House overall | Numerical | 0 | 14714.57 | W |

## 1.4 Question to be addressed

# Formulation of the mathematical model

## 2.1 LSTM model

Diagram

Description automatically generatedLSTM is specifically made to prevent the long-term dependency problem. The cord idea of LSTM is the cell state and the structure of gate. The cell state is the path of information transmission, so the information can be passed in a serial connection. A control signal of the memory gate controls whether the gate should retain the information. In the implementation, it is normally multiplied by 0 or 1 to choose to retain or forget. Even information from the earlier time steps can be carried to cells at later time steps, which overcomes the effect of short-term memory. Fig. shows the procedure of forward propagation of LSTM. Fig. (a), (b), (c), and (d) depicts the calculation process of the forget gate, input gate, cell state, and output gate, respectively.

**Fig. 1.** Procedure of forward propagation of LSTM.

# Solution of the problem

# Interpretation of results

# Critique of the model

# Appendices with supporting materials

**Reference**

[1] G. Pinto, D. Deltetto, A. Capozzoli, Data-driven district energy management with surrogate models and deep reinforcement learning, Applied Energy. 304 (2021) 117642. https://doi.org/10.1016/j.apenergy.2021.117642.

[2] Y. Peng, A. Rysanek, Z. Nagy, A. Schlüter, Occupancy learning-based demand-driven cooling control for office spaces, Building and Environment. 122 (2017) 145–160. https://doi.org/10.1016/j.buildenv.2017.06.010.

[3] X.J. Luo, L.O. Oyedele, Forecasting building energy consumption: Adaptive long-short term memory neural networks driven by genetic algorithm, Advanced Engineering Informatics. 50 (2021) 101357. https://doi.org/10.1016/j.aei.2021.101357.

[4] I. Karijadi, S.-Y. Chou, A hybrid RF-LSTM based on CEEMDAN for improving the accuracy of building energy consumption prediction, Energy and Buildings. 259 (2022) 111908. https://doi.org/10.1016/j.enbuild.2022.111908.