Memorial University of Newfoundland

Department of Electrical and Computer Engineering

Electric Load Forecasting Using Deep Neural Networks

Presented by:
Zain Ahmed

Supervisors:
Dr. Mohsin Jamil and Dr Ashraf Ali Khan

Date: 12 September 2024

Presentation Outline

- **Introduction**
  - Background
  - Literature Review
  - Research Objectives

- **Short-Term Campus Load Forecasting using CNN-based Encoder-Decoder Network with Attention**
  - Data Description
  - Network Architecture
  - Model Pipeline
  - Results and Analysis

- **A Novel Multi-Task Learning Based Approach to Multi-Energy System Load Forecasting**
  - Introduction
  - Data Description
  - Proposed training model
  - Results and Analysis

- **Future Direction**

- **Pulications**

Background

- Load Forecasting is an important tool for power service provider.
- It aids the daily operation and maintenance activities.
- It helps identify peak hours and prevent load shedding.
- It helps in planning maintenance activities.
- Accurate long term forecasts can affect policies and future direction of the company.
- A one percent increase in forecasting error is associated with a $10 million increase in operating costs.

Literature Review

- A load forecasting problem can be formulated as a time series forecasting problem and divided into different categories based on:

  1. The Output Horizon Length

  2. Nature of the Algorithm Used

  3. Number of input variables employed

.

Literature Review

- Based on the Output Horizon Length the load forecasting can be divided into:

1. Short Term Load Forecasting:

   Output Horizon is between one hour to one week

2. Medium-Term Load Forecasting:

   Output Horizon is between one week and one year

3. Long-Term Load Forecasting:

   Output horizon is between longer than one year

Literature Review

- Based on the nature of algorithm used the load forecasting can be divided into:

  1. Statistical Load Forecasting Techniques

  2. Artificial Intelligence-based Techniques

  3. Hybrid Techniques

- Based on the number of input variables employed load forecasting can be divided into:

  1. Univariate Load Forecasting

  2. Multivariate Load Forecasting

Literature Review

- Major papers which were instrumental in the research include:

.

Reviewed Paper	Description
M. Madhukumar et al. (2022)	Regression Model-Based Short-Term Load Forecasting for University Campus Load
M. Aouad et al. (2022)	A CNN-Sequence-to-Sequence network with attention for residential short-term load forecasting
Y. Guo et al. (2022)	BiLSTM Multitask Learning-Based Combined Load Forecasting Considering the Loads Coupling Relationship for Multienergy System

Research Objectives

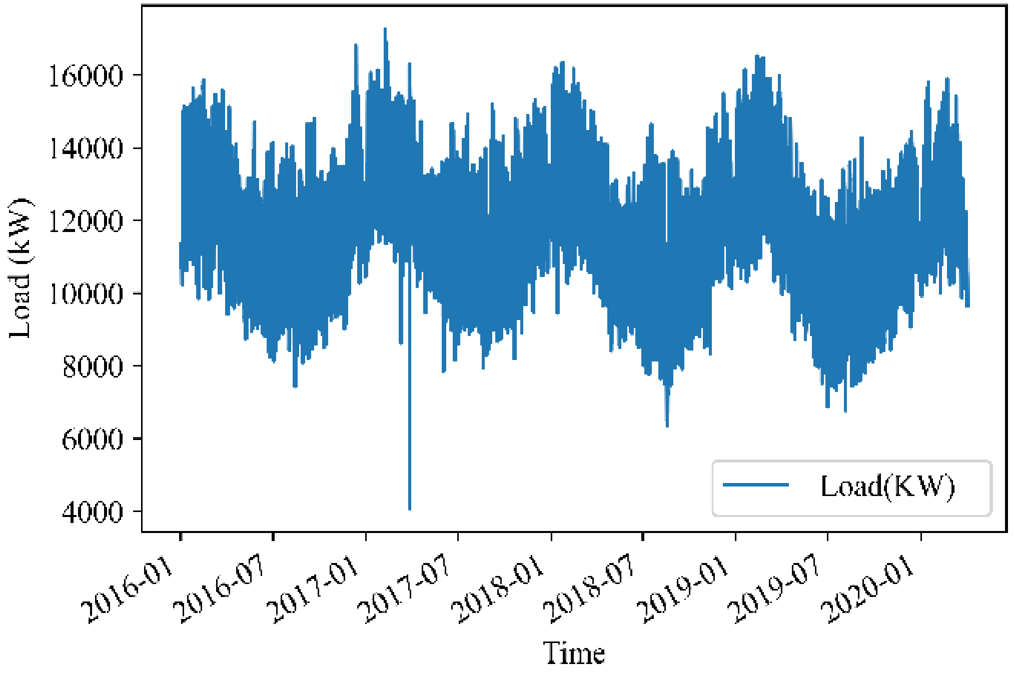
- The first part of the research is focused on a novel architecture for campus load forecasting using an attention-based sequence-to-sequence model.
- The second part if focused on a novel multi-task learning based approach for load forecasting of multi-energy systems.

.

Short-Term Campus Load Forecasting using CNN-based Encoder-Decoder Network with Attention

Data Description

- MUN Campus Load Data from 2016 to 2020 was used



MUN electric load

Parameter	Value
Data start date	January 2nd 2016
Data end date	2020st 31March
Data interval	1 hour
Total data points	37221
Features	8

Load Data Features

Data Description

- Seven metrological features were employed
- These include:

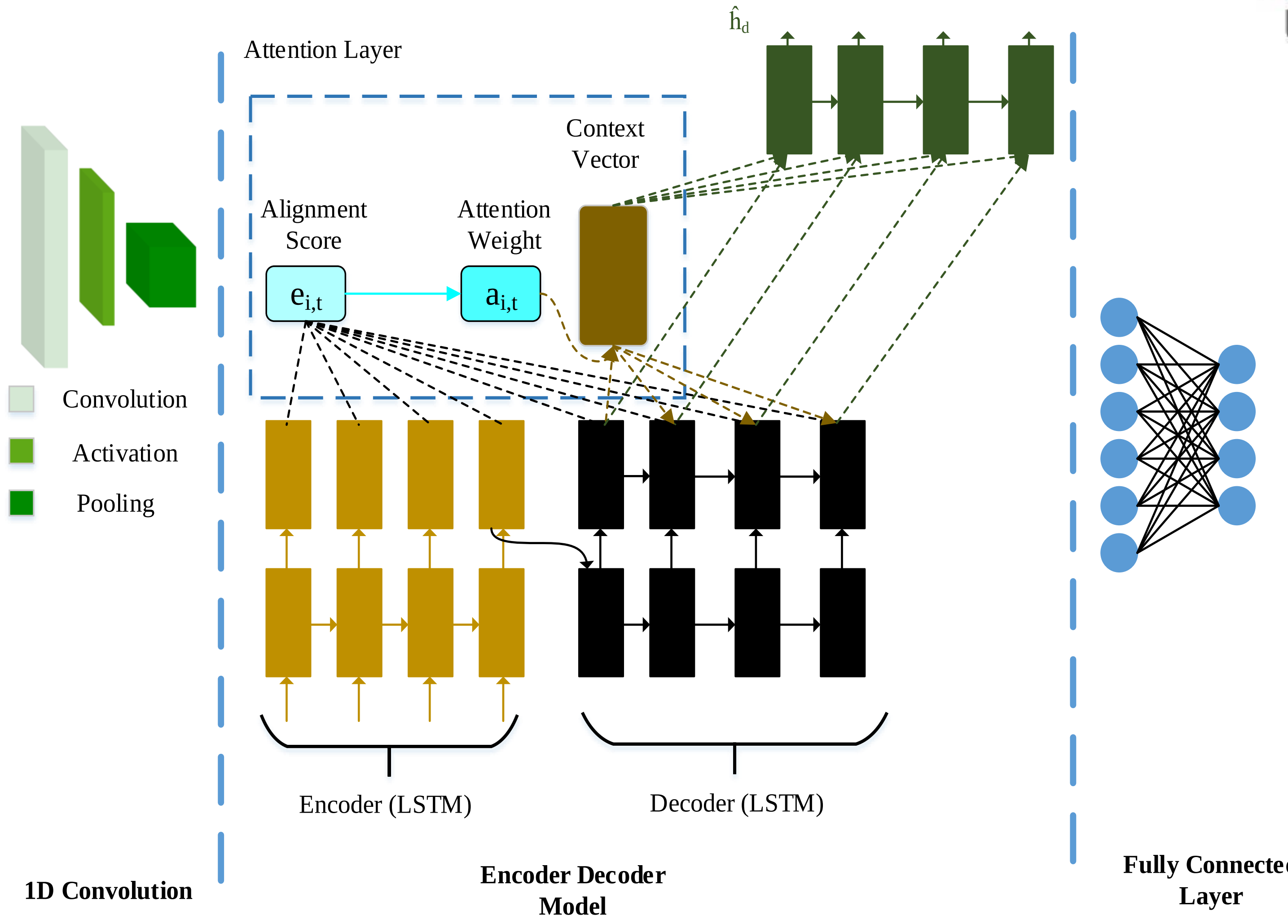
1. Dry Bulb Temperature
2. Dew Point
3. Relative Humidity
4. Wind Direction
5. Wind Speed
6. Visibility
7. Atmospheric Pressure

Network Architecture

- The proposed network architecture consists of the following major components:

1. One Dimensional Convolutional Neural Network to capture relations among different features
2. An Long Short Term Memory based Sequence to Sequence Model to capture temporal features
3. Fully Connected Layer to make predictions

Network Architecture



Proposed Network Architecture

Model Pipeline

- The Model Pipeline consists of the following major components:

1. Data Cleaning
    - Removal of outliers
    - Removal of NA values from the dataset
2. Data Preprocessing:
    - Data standardization
    - Data Windowing
    - Data Division

Model Pipeline

3. Model Training

- Data is fed to the network architecture

- Hyperparamater Optimization

Hyperparameters	Values
Epochs	100
Loss Function	Huber Loss
Batch Size	128
Optimizer	Adams Optimizer
Early Stopping Patience	20
Early Stopping Parameter	Validation Loss

Results and Analysis

- For analyzing the results, we will briefly describe the evaluation metrics
- Mean absolute error (MAE) is obtained by taking the mean of absolute difference between actual and predicted values. Lower MAE demonstrates better algorithm performance.
- Mean Squared Error (MSE) is the average of the squared difference between the predicted values and ground truth. Lower MSE demonstrates better algorithm performance.

Results and Analysis

- It is also referred to as the coefficient of determination. It can have a maximum value of 1 which would indicate that the model has predicted every ground truth value correctly. It can also have negative values with no limit since an estimator can be arbitrarily worse.
- Mean Absolute Percentage Error (MAPE) is an evaluation metric that is defined below. G_i here is the ground value while P_i represents the predicted values :

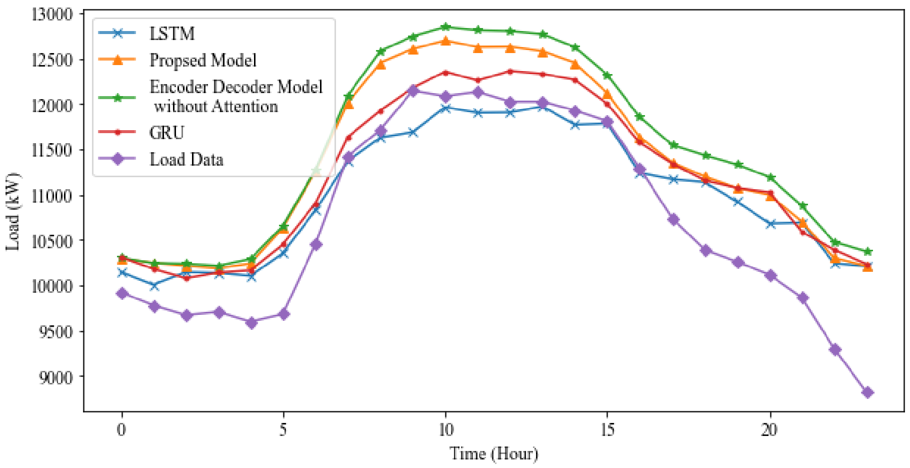
$$MAPE = \frac{100}{n} \sum_{i=1}^{N} \left| \frac{G_i - P_i}{G_i} \right|$$

Results and Analysis

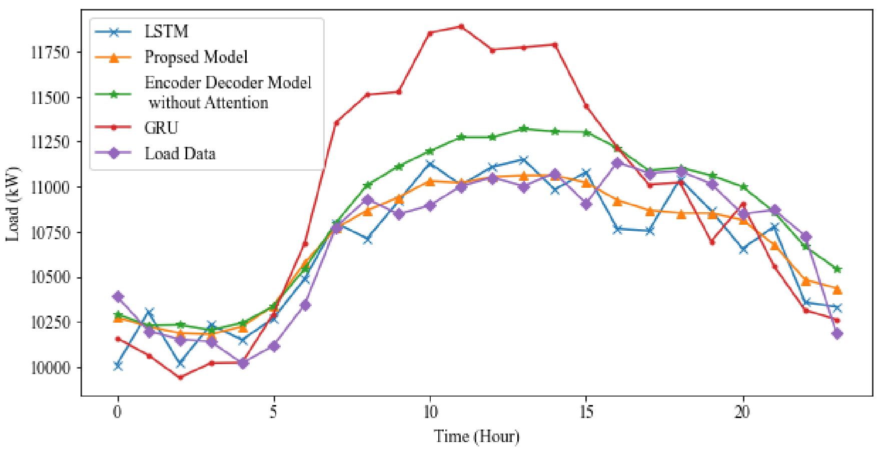
Algorithm/Metric	MAE (kW)	R2 Score	MSE (kW2)	MAPE (% age)
LSTM	442.94	0.77	340026.33	3.64
GRU	493.72	0.739	386019.62	4.05
1D CNN + Encoder Decoder	423.326	0.8011	294162.21	3.51
RQ-GPR [12]	450.21	0.71	401264.54	4.98
Sequence to Sequence + Luong Attention [14]	410.11	0.795	298765.45	3.40
Proposed Network Architecture	407.308	0.805	287346.33	3.37

Results and Analysis

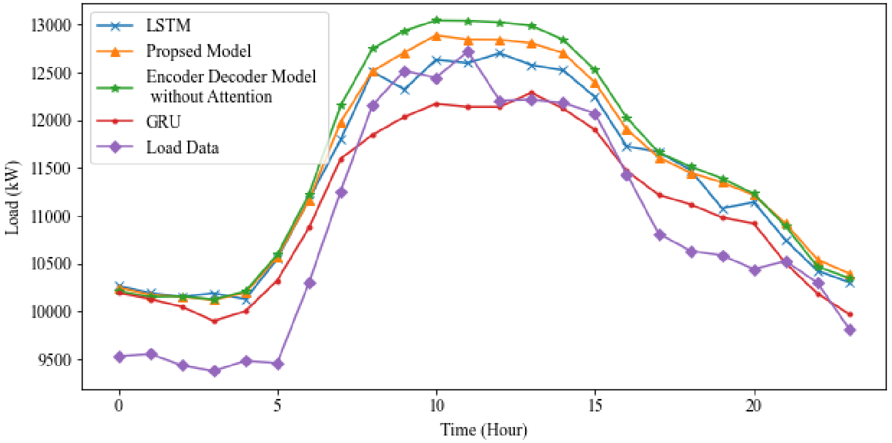
Case 1



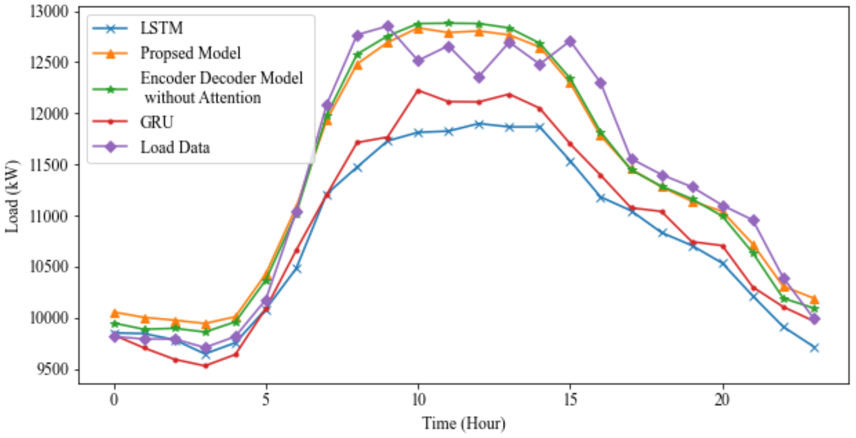
Case 2



Case 3



Case 4

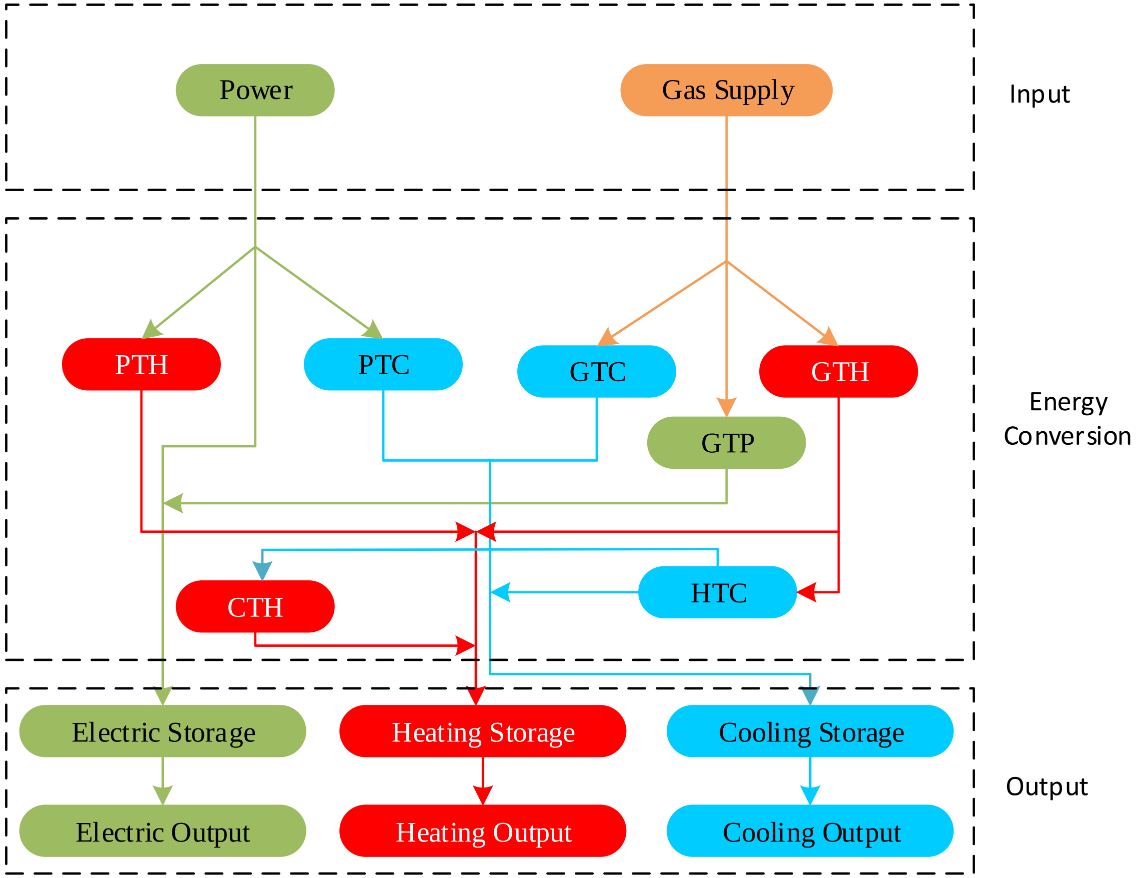


A Novel Multi-Task Learning Based Approach to Multi-Energy System Load Forecasting

Introduction

- A Multi-Energy System (MES) is an integrated approach to managing and supplying different types of energy carriers within a single coordinated unit.
- MES utilizes the transfer among different forms of energy to offer flexibly, improved efficiency, environmental performance and sustainability
- The proposed method uses a novel multi-task learning based approach for multi-energy systems.

Introduction



Generic Multi-Energy System

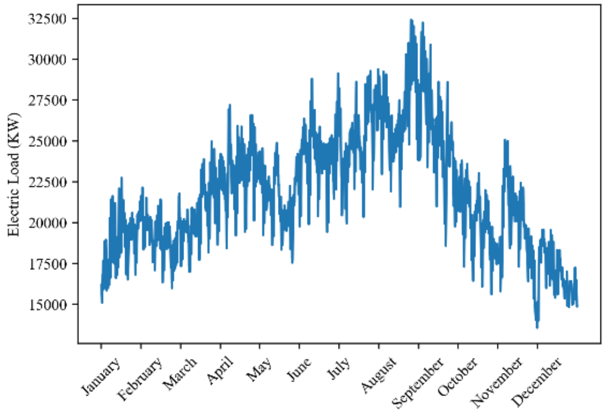
Data Description

- The data is taken from University of Austin Tempe Campus
- It consists of heating, cooling and electric loads

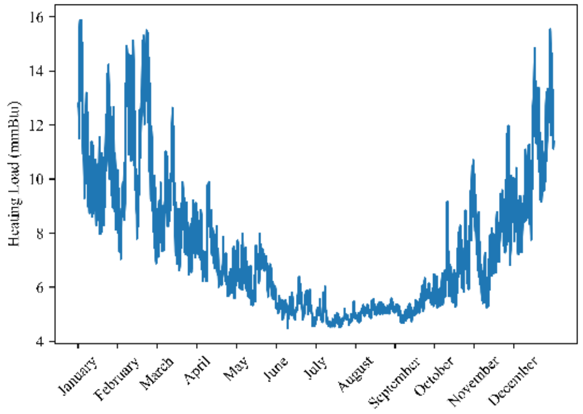
Parameter	Value
Data start date	1st January 2016
Data end date	31st December 2019
Data interval	1 hour
Number of observations	35063
Number of metrological features	6
Number of temporal features	21

Salient Data Features

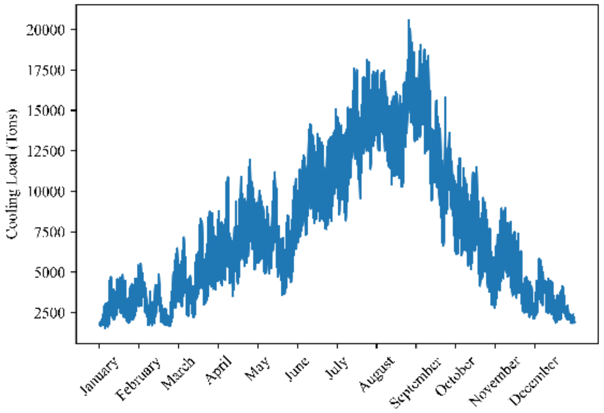
Data Description



Electric Load Data from 2019

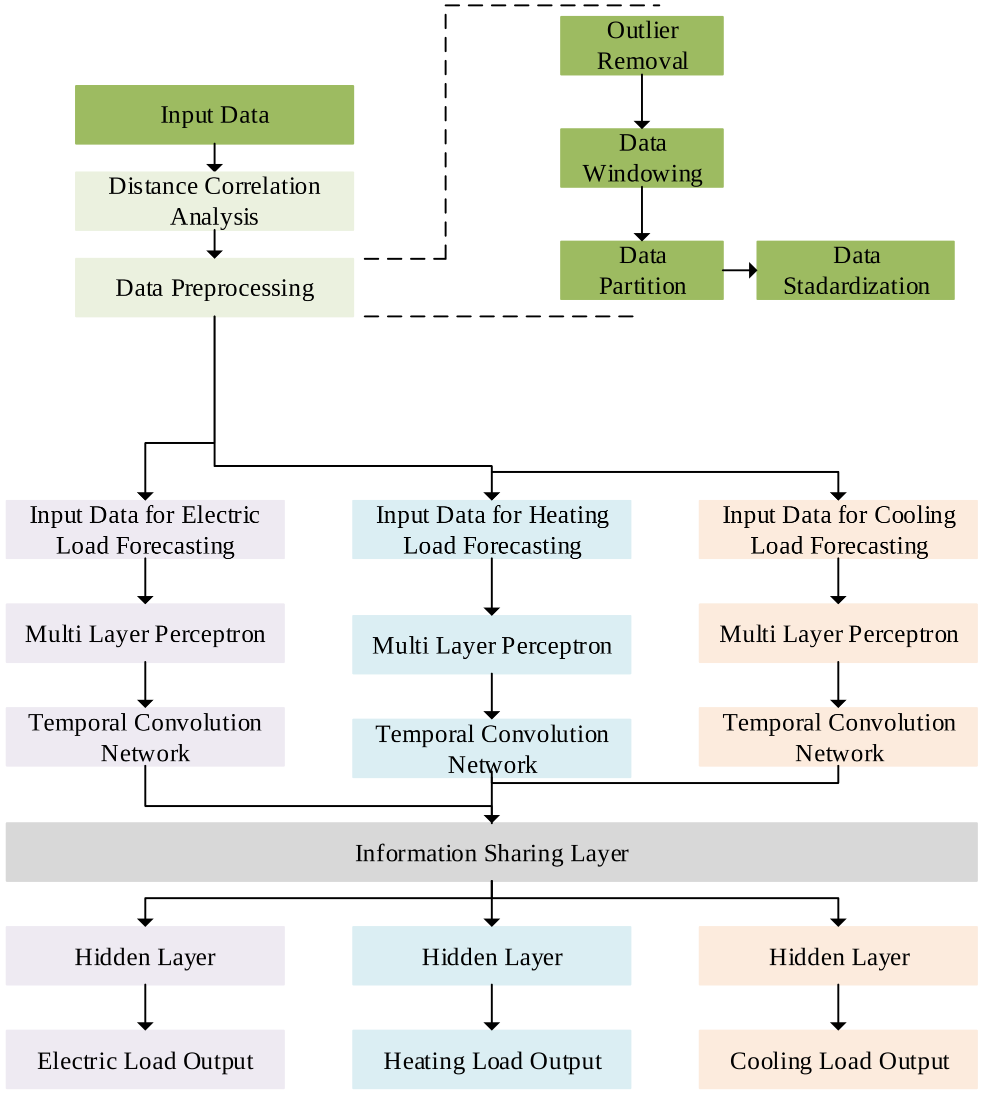


Heating Load Data from 2019



Cooling Load Data from 2019

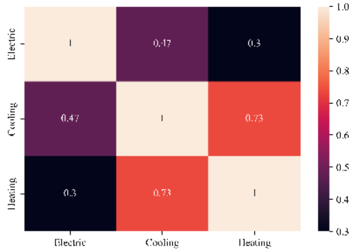
Proposed Training Model



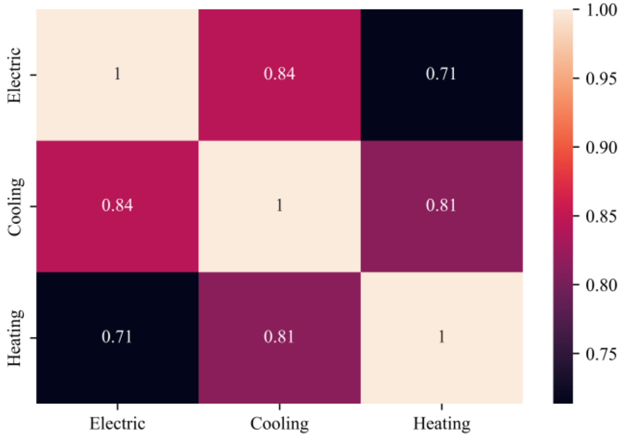
Proposed D-TCNet Architecture

Proposed Training Model

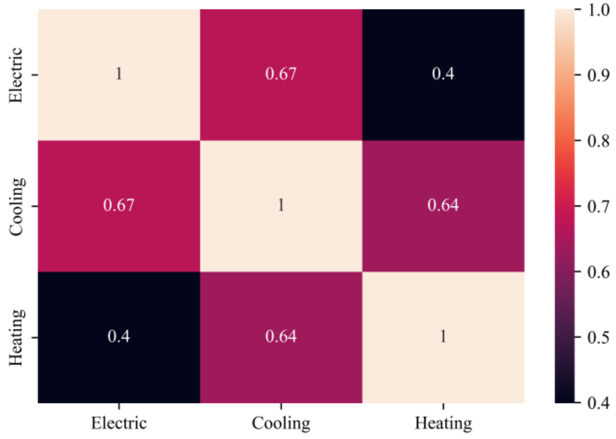
- First off, data is divided into four seasons (summer, winter, spring and autumn).
- Distance Correlation Analysis is performed between electric load, heating load and cooling load for each season.
- Distance Correlation or Distance Covariance is a metric used to measure the degree of dependence or interrelationships between two random variables or vectors.
- Based on the analysis, the load variables with high interrelationships are used as input metrics while others are discarded.



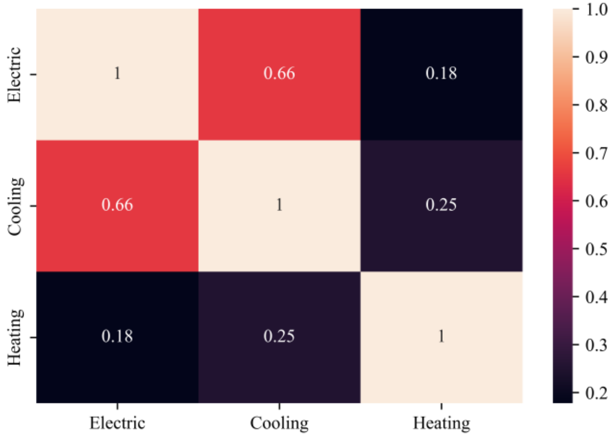
osed Training Model



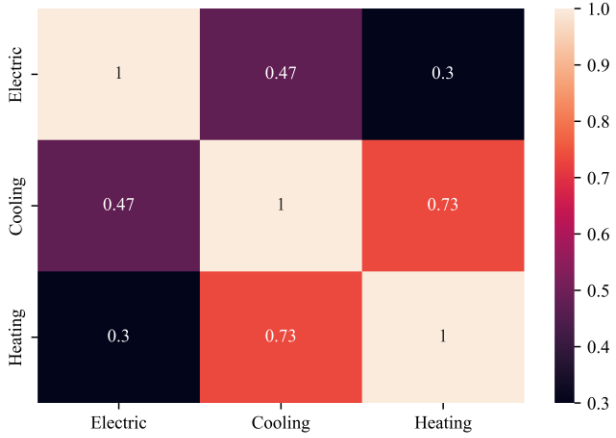
Autumn Distance Correlation



Spring Distance Correlation



Summer Distance Correlation



Winter Distance Correlation

Distance Correlation heat maps

Proposed Training Model

- Data preprocessing starts with outlier removal using interquartile range (IQR) method. Any data outside the range [Q1 – 3 * IQR, Q3 + 2 * IQR] was filled using forward filling method.
- Data windowing is done to create proper input-output pairs. Input window size of 168 is employed to provide the algorithm with hourly data profile of the previous week while output window size is set to 1.

Proposed Training Model

- Next data portioning is carried out. The classical 80/10/10 method is used here. It means that 80% of the dataset is used for training while 10 percent for validation and the remaining 10% as test set.
- Final data preprocessing step is data standardization to ensure fast convergence times.
- Below μ_i is the mean while σ_i is the standard deviation for ith feature while $\hat{x}_i$ is the standardized feature

$$\hat{x}_i = \frac{x_i - \mu_i}{\sigma_i}$$

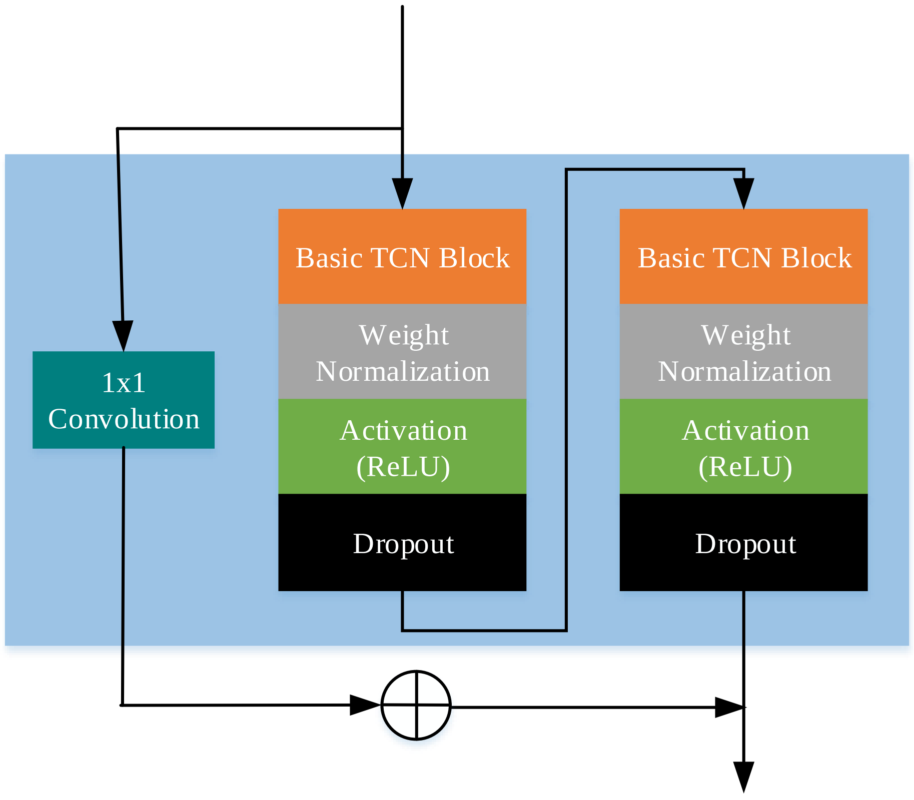
Proposed Training Model

- Data is then fed to the Deep Temporal Convolutional Network (D-TCN).
- First layer is Multi Perceptron Layer for each of the three data streams (heating, electric and cooling) embed multiple features.
- Next is the Temporal Convolutional Block to find temporal patterns in the data.
- Data from these three streams is concatenated in the feature sharing layer and finally predictions are made from there.

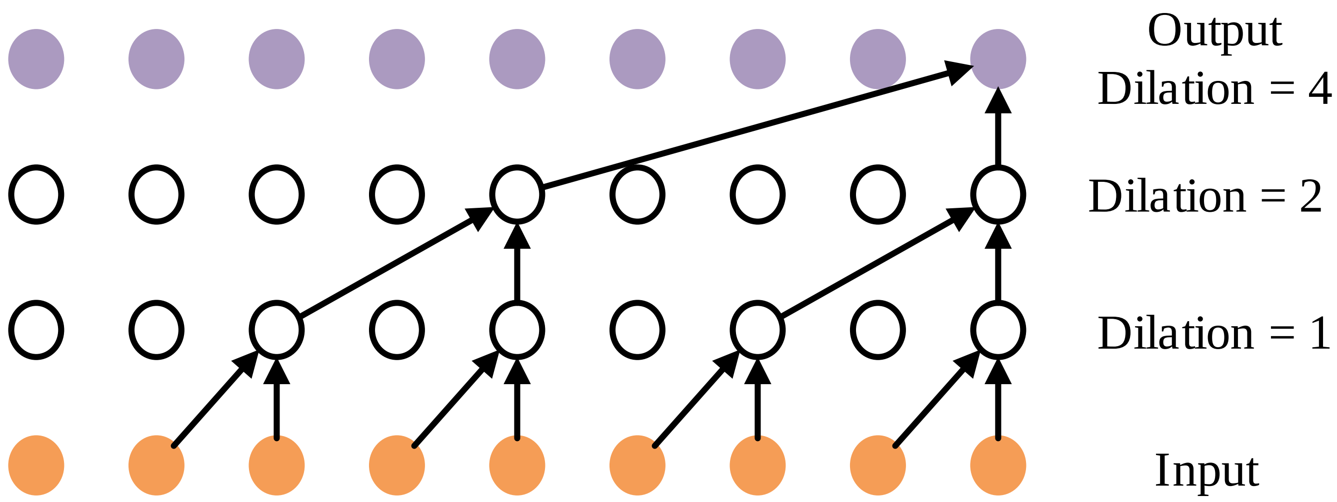
Proposed Training Model

- Temporal Convolutional Network (TCN) are convolutional neural networks used for handling temporal data.
- TCN mainly consists of casual and dilated convolutions.
- Casual convolutions ensure that model only learns from past time steps.
- Dilated convolutions improve the receptive field of the network.

Proposed Training Model



A complete residual block



Simple TCN block

Proposed Training Model

Parameter	Value
MLP Embedding dimension	8
TCN kernel size	8
TCN Dropout	0.2
TCN activation function	Leaky ReLU
Number of TCN Blocks	3

Network Parameters

Parameter		Value
Max epochs		200
Initial Learning Rate		0.01
Learning rate reduction on plateau	Reduction Patience	10
	Reduction factor	0.5

Training Routine Parameters

Results and Analysis

Results for Autumn

Algorithm	Load Type	Evaluation Metric
		MAPE (%)
MIC + Bi-LSTM [10]	Electric	2.2
	Heating	**4.1**
	Cooling	7.4
Single-Task Approach	Electric	2.4
	Heating	4.4
	Cooling	7.9
Multi-Task Approach	Electric	**1.9**
	Heating	4.2
	Cooling	**6.9**

Results for Spring

Algorithm	Load Type	Evaluation Metric
		MAPE (%)
MIC + Bi-LSTM [10]	Electric	2.1
	Heating	3.7
	Cooling	3.5
Single-Task Approach	Electric	2.2
	Heating	3.8
	Cooling	3.5
Multi-Task Approach	Electric	**2.1**
	Heating	**3.5**
	Cooling	**3.1**

Results and Analysis

Results for Summer

Algorithm	Load Type	Evaluation Metrics
		MAPE (%)
MIC + Bi-LSTM [10]	Electric	2.4
	Heating	2.1
	Cooling	2.2
Single-Task Approach	Electric	2.5
	Heating	2.3
	Cooling	2.1
Multi-Task Approach	Electric	**2.1**
	Heating	**1.5**
	Cooling	**2.0**

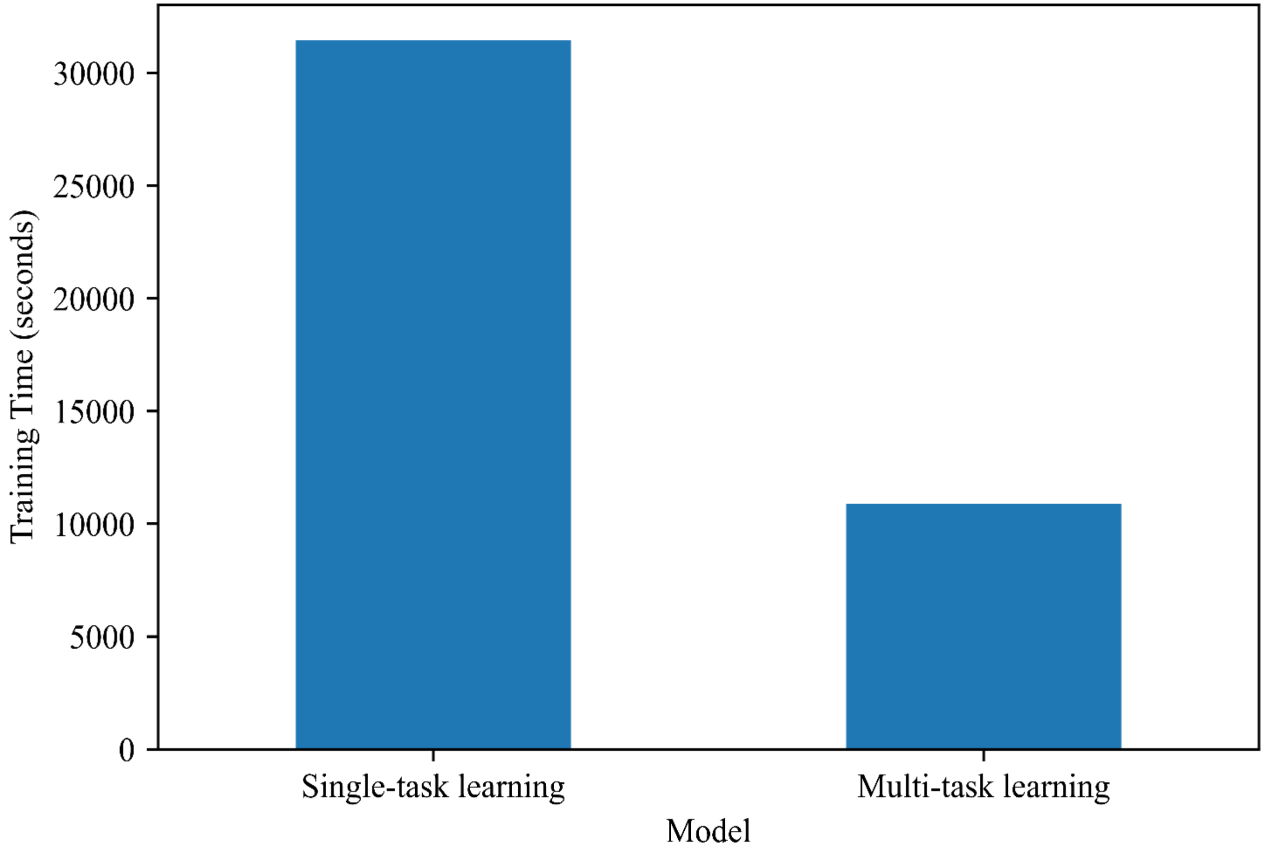
Results for Winter

Algorithm	Load Type	Evaluation Metrics	
		MAPE (%)	RMSE (KW)
MTL-LSSM	Electric	1.9	
	Heating	3.5	
	Cooling	3.9	
Single-Task Approach	Electric	2.6	
	Heating	3.6	
	Cooling	5.2	
Multi-Task Approach	Electric	**1.4**	
	Heating	**2.7**	
	Cooling	**2.1**	

Results and Analysis

- Proposed algorithm shows best performance in all four seasons compared to other state of the art algorithms
- The multi-task approach performs better than single-task approach which shows that the feature sharing layer improves model performance.
- Multi-task learning approach also helps improve the training times. This is shown on the next graph.

Results and Analysis



Training time comparison

Future Direction

- Introduction of bi-directional encoder decoder instead of the standard encoder decoder to investigate its performance on multiple electric load datasets.
- Performing medium and long term load forecasting instead of short term load forecasting by leveraging long range temporal dependencies of the attention mechanism
- For multi-task learning based approach, investigation of encoder-decoder model instead of TCN block for multi-energy load forecasting

List of Publications

- Z. Ahmed, M. Jamil, A.A. Khan, “Short-Term Campus Load Forecasting Using CNN-Based Encoder–Decoder Network with Attention”. *Energies* **2024**, *17*, 4457. https://doi.org/10.3390/en17174457
- Z. Ahmed and M. Jamil, "Campus Electric Load Forecasting Using Recurrent Neural Networks," *IEEE12th International Conference on Smart Grid (icSmartGrid)*, Setubal, Portugal, 2024, pp. 412-417, doi: 10.1109/icSmartGrid61824.2024.10578153.
- Z. Ahmed, M. Jamil, A.A. Khan, “A Novel Multitask Learning-Based Approach to Multi-Energy System Load Forecasting,” Revision submitted to IEEE Open Journal of Power and Energy.


ANY
QUESTIONS?