Energy Consumption Prediction

2024-2025

Cristiano Nicolau, 108536
Nelson Loureiro, 124026

Table of contents

Introduction

Context:

Development of ML and DL models aimed at predicting energy consumption in buildings of various types.

Objective:

Using comprehensive data on building characteristics, weather conditions, and historical energy usage:

- Predict the energy consumption of a building for the upcoming moments.
- Identify and model consumption patterns over time.

Motivation:

- Estimate energy consumption in the near future, enabling better energy management.
- Optimize the planning of energy infrastructure and the efficient allocation of resources in smart grids, contributing to sustainability and innovation in urban energy systems.

State of Art

Fast Prediction Models
Prioritize speed and efficiency, making them ideal for real-time energy consumption forecasting. Models like Feedforward Neural Networks (FNN) and Random Forests (RF) are used. These models are often used in real-time energy monitoring systems.

Multistage Models
Handling large-scale data, such as country-level or regional energy consumption forecasts, multistage models are ideal. These models break the problem into several phases, with each phase focusing on different aspects of the data or forecasting process.

Highly Accurate Models
These models are designed for high precision, often focusing on individual buildings or more detailed energy usage patterns. They can accommodate complex relationships and temporal dependencies, which are critical for accurate short-term and long-term forecasts. Models like Long Short-Term Memory (LSTM) networks and Bidirectional LSTM (Bi-LSTM) are particularly effective in this area. These architectures capture patterns in sequential data, such as the fluctuations in energy consumption throughout the day or across seasons.

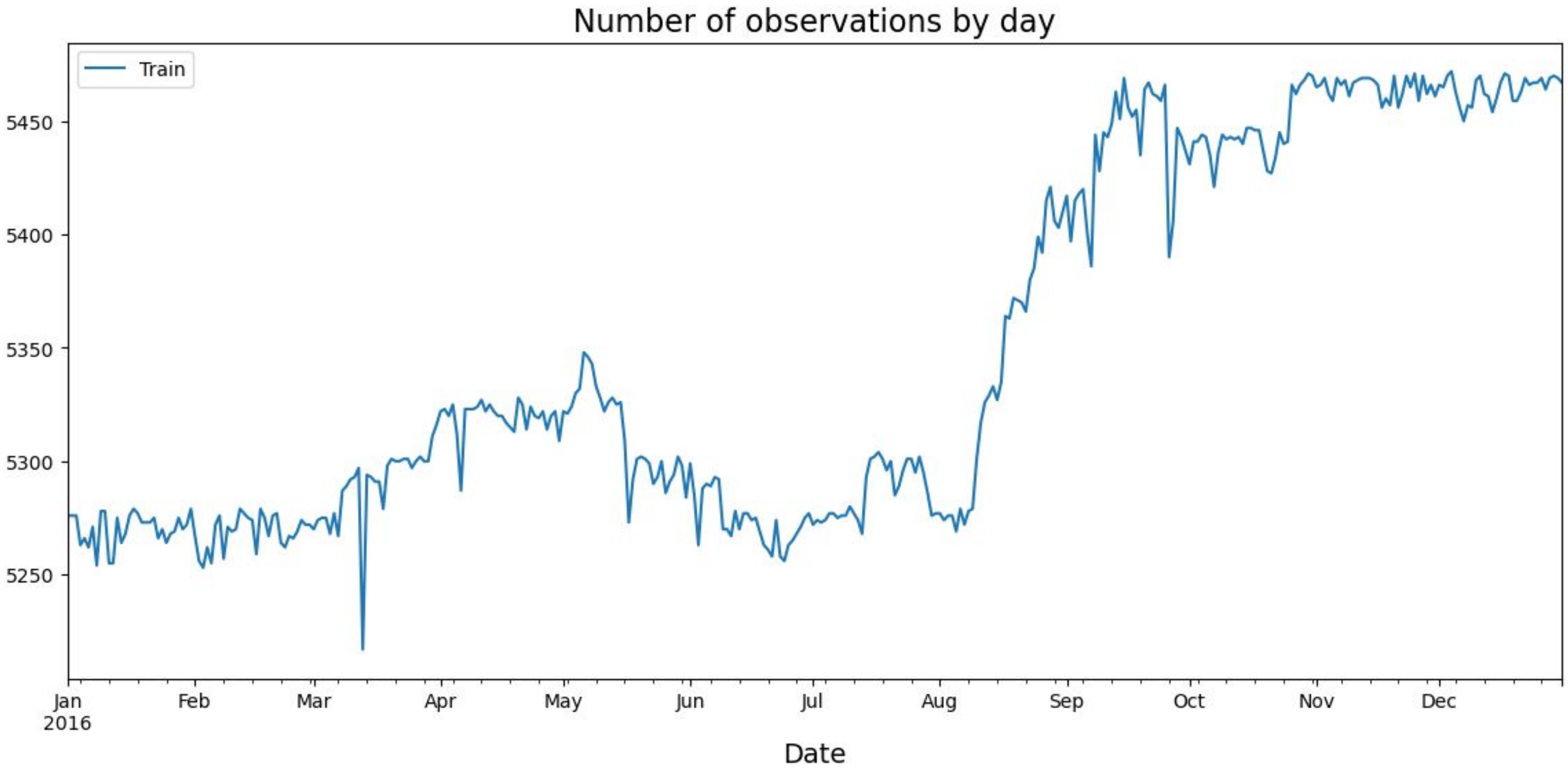
Data

- Sourced from the **ASHRAE - Great Energy Predictor III** competition on Kaggle (2019)
- **Energy usage measurements across multiple energy types** (electricity, chilled water, steam, hot water) from over 1,400 commercial and institutional buildings
- `train.csv`: ~20.5 million records with energy readings, meter id, building id, site id
- `weather_train.csv`: ~1.4 million records
- Building features: area (m^2), year built, number of floors, primary use, etc.
- Weather data: air temperature, humidity, pressure, wind speed and direction, among others
- Data from **16 different locations** over the course of **one full year**
- Presence of **extreme values (outliers)** and **missing data** in several columns

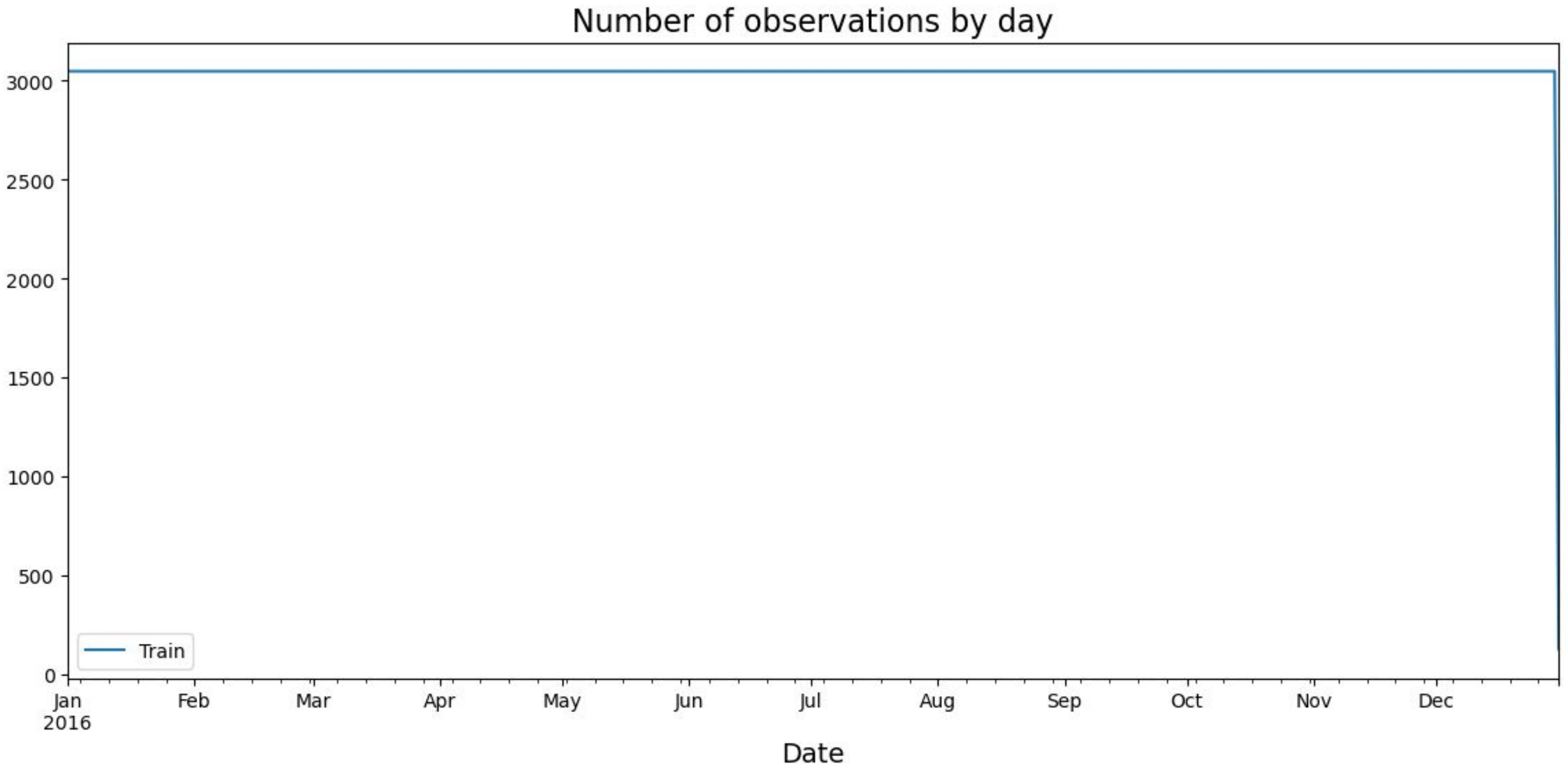
Data - Pre Processing

- Load data and merge datasets using `building_id`, `site_id`, and `timestamp` to combine building and weather info
- Filtered for **electricity consumption only** (`meter = 0`), reducing dataset size
- Dropped Buildings with missing `year_built` and `floor_count`, reducing data from **12M to 2M rows** to improve reliability
- Filled missing weather values with **site-wise medians**; remaining nulls filled with **global medians**
- Removed low-correlation columns like `wind_direction`, `wind_speed`, and `sea_level_pressure`

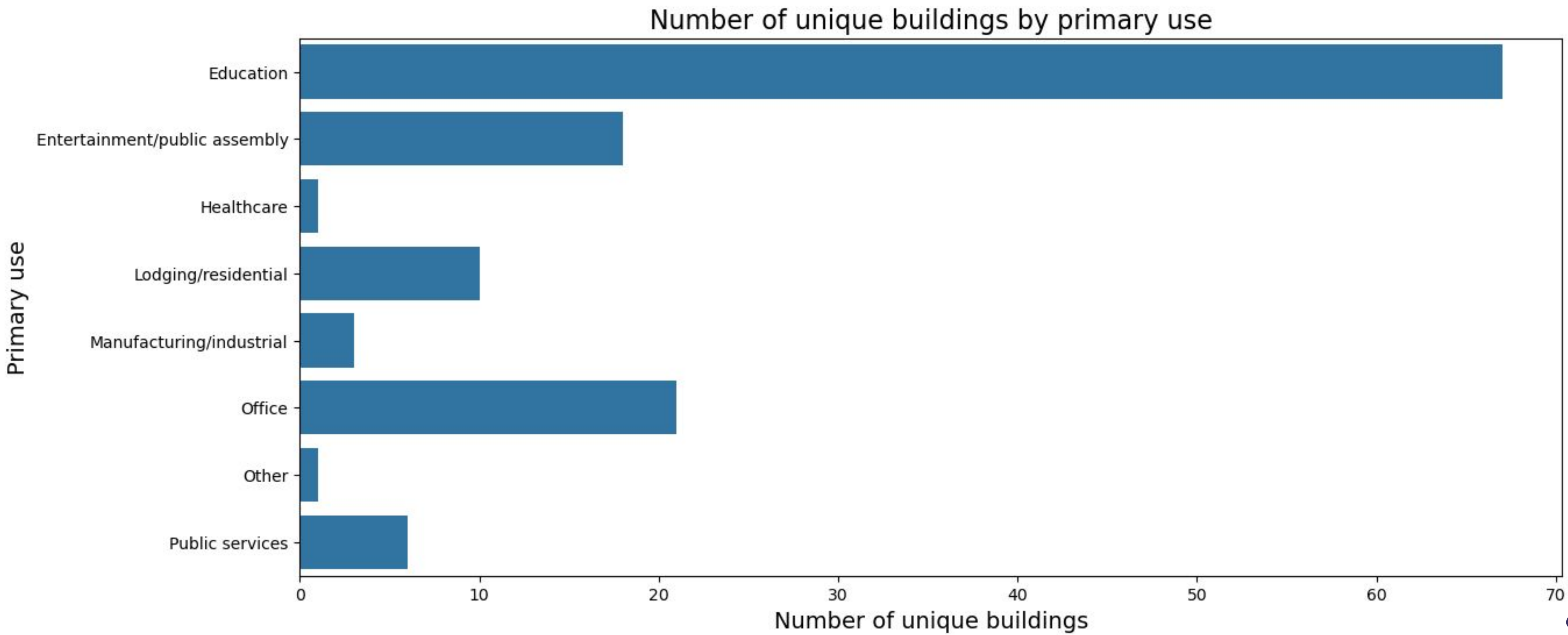
Data Processing - Clean Data



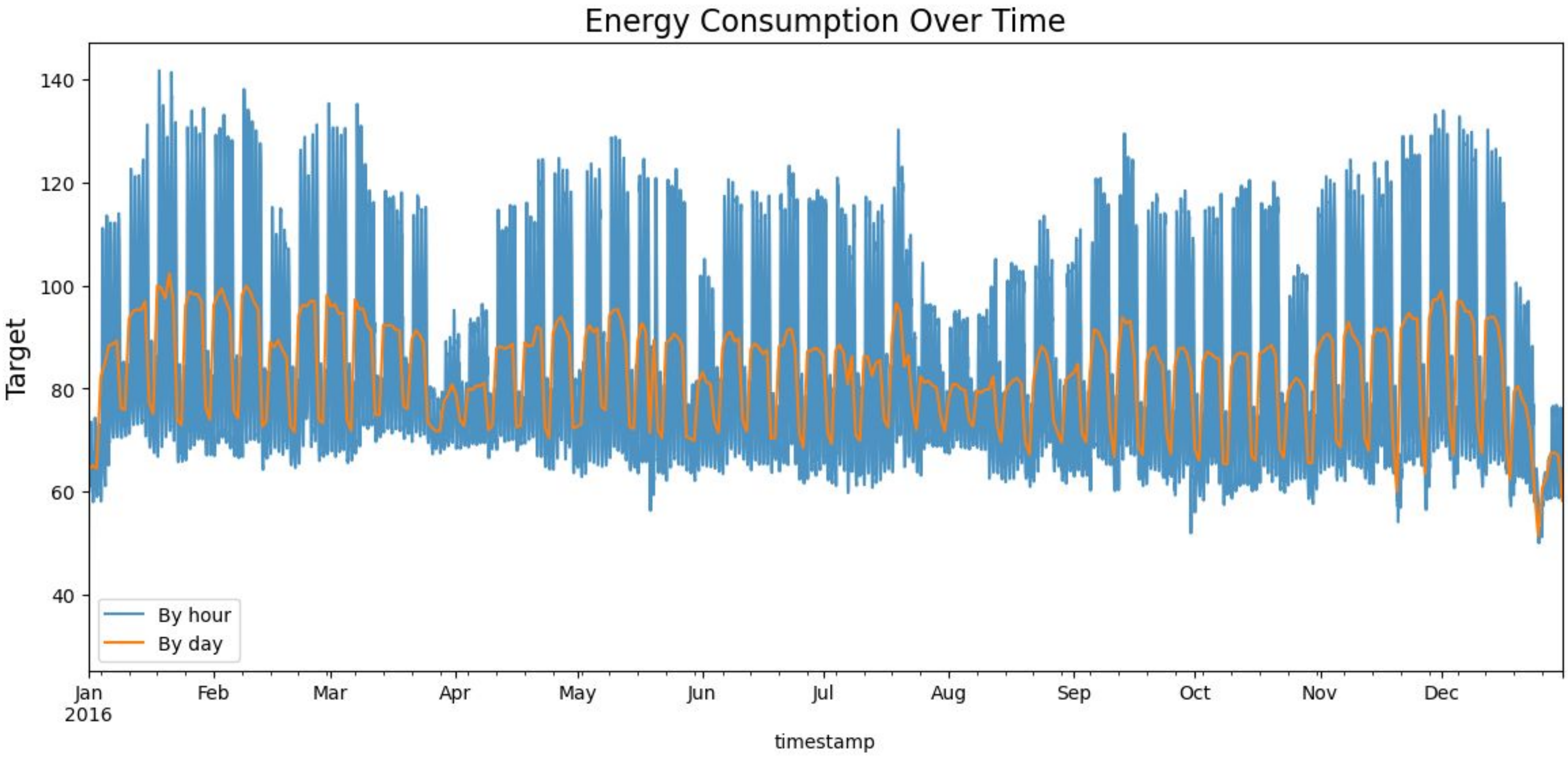
Data Processing – Clean Data



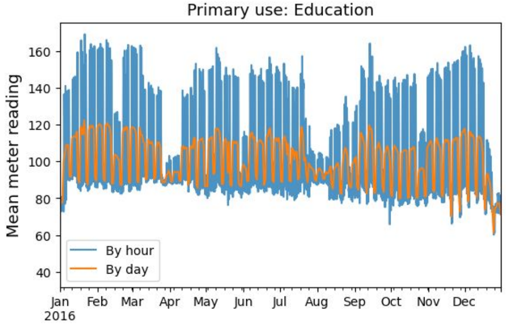
Data Processing - Clean Data

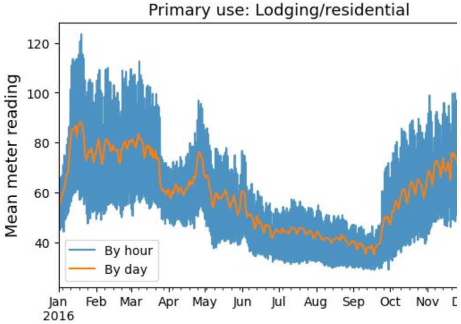


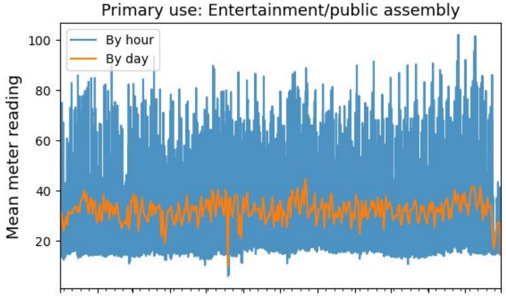
Data Processing - Data Visualization

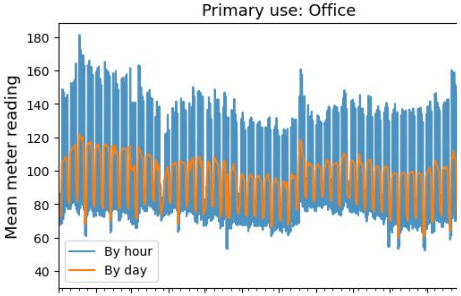


Data Processing - Data Visualization



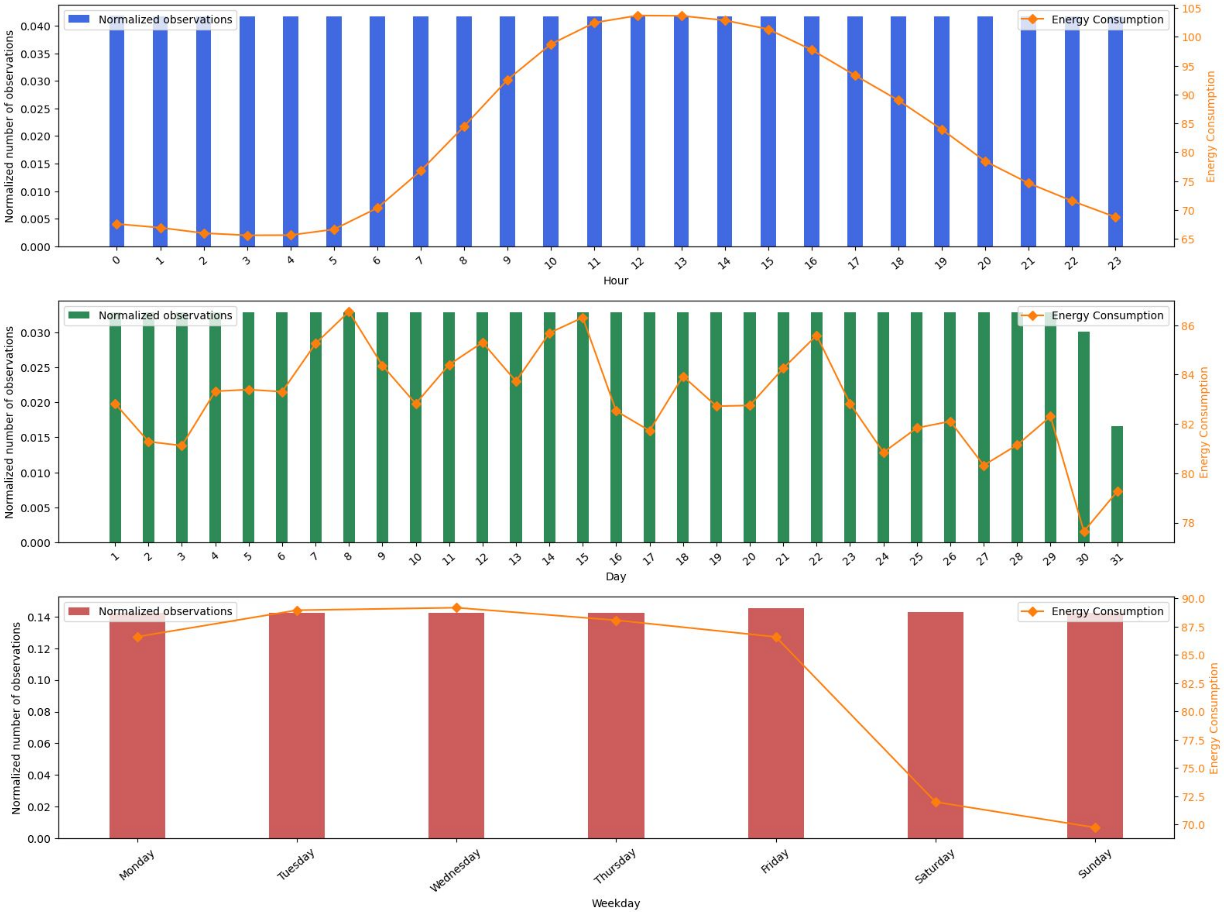




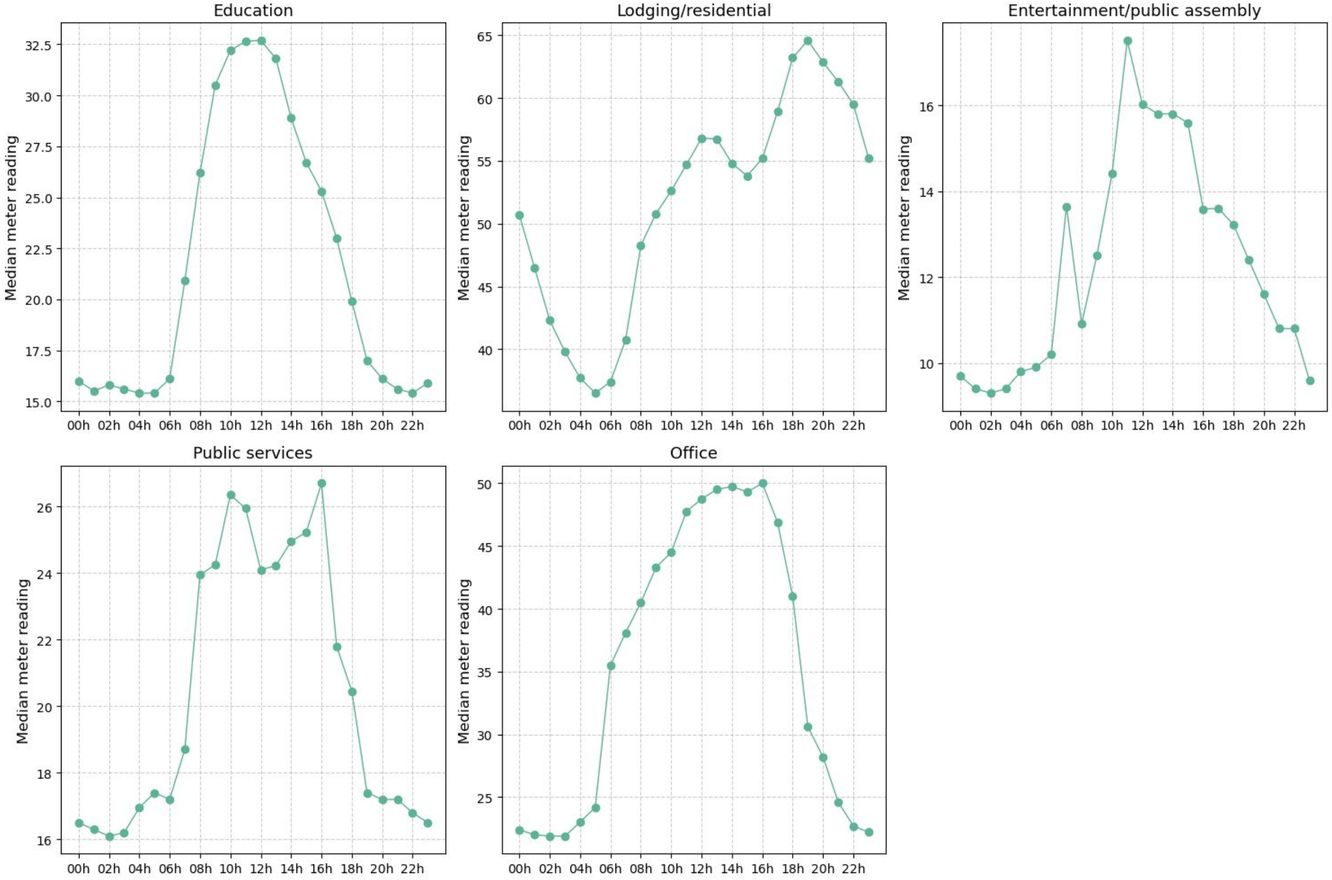


primary_use	number_of_buildings
Education	67
Entertainment/public assembly	18
Lodging/residential	10
Office	21
	116

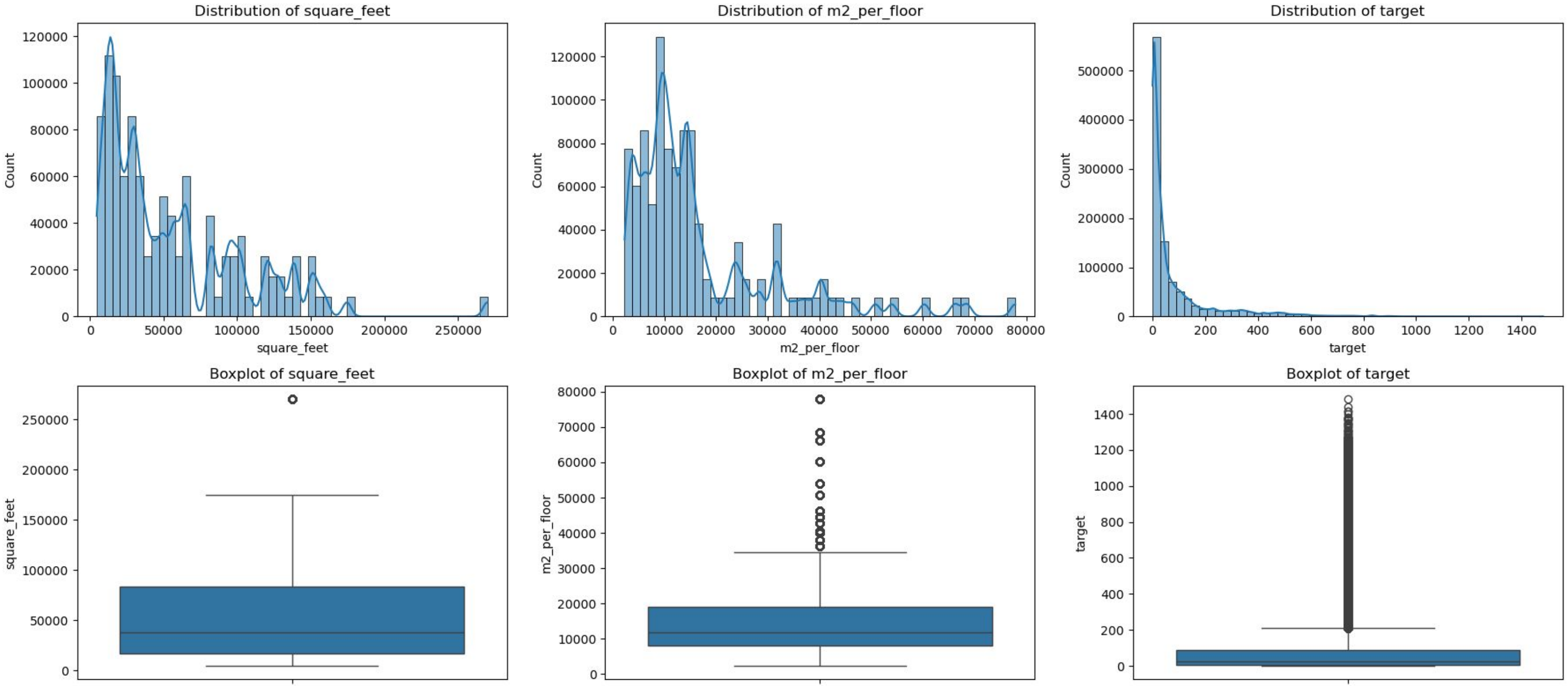
Data Processing - Data Vis



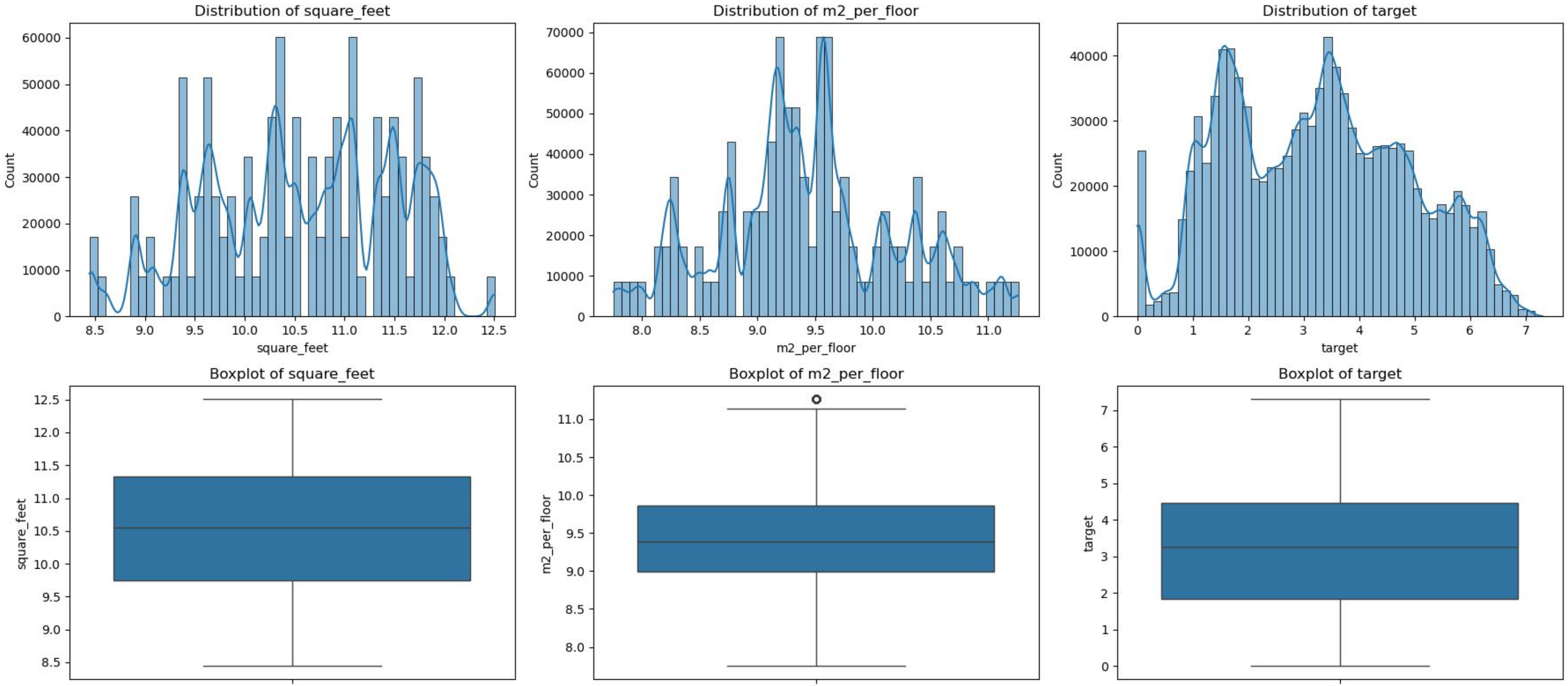
Data Processing - Data Vis



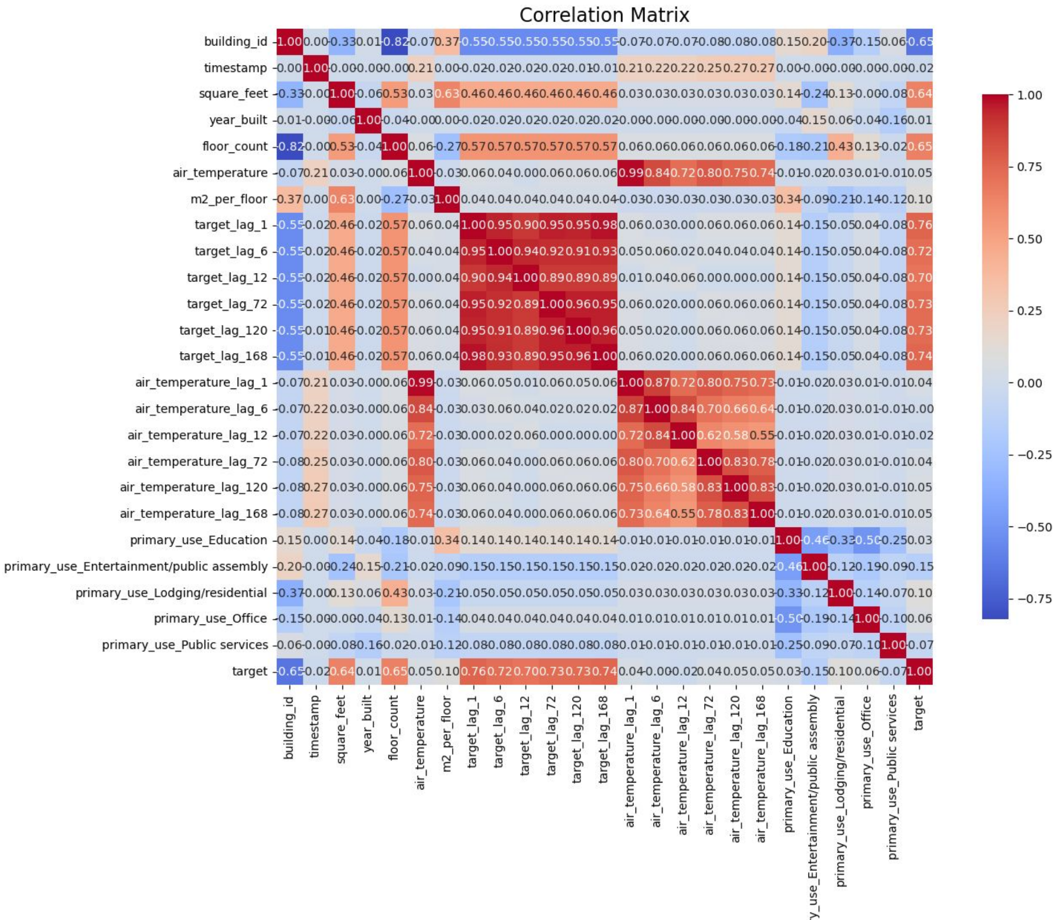
Data Processing - Data



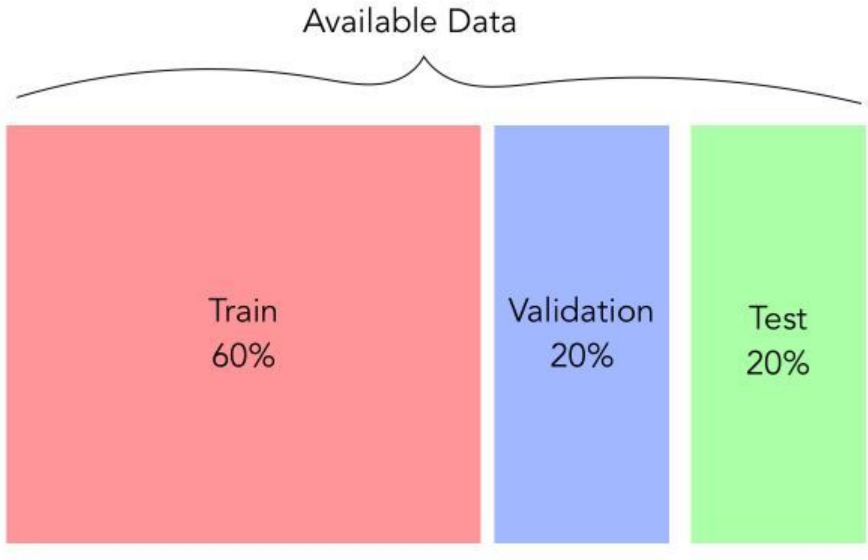
Data Processing - Data Visualization



Data Processing - Data Visual



METHODOLOGY



RF - Model

RF Hyperparameters

Hyperparameter	Default Value	Tuned Value
n_estimators	100	500
min_samples_split	2	2
min_samples_leaf	1	2
max_features	'auto'	'log2'
max_depth	None	None

Randomized Hyperparameter Search

Hyperparameter	Search Space
n_estimators	[100, 200, 300, 400, 500]
max_depth	[None, 10, 20, 30, 40, 50]
min_samples_split	[2, 5, 10]
min_samples_leaf	[1, 2, 4]
max_features	['auto', 'sqrt', 'log2']

RF Results

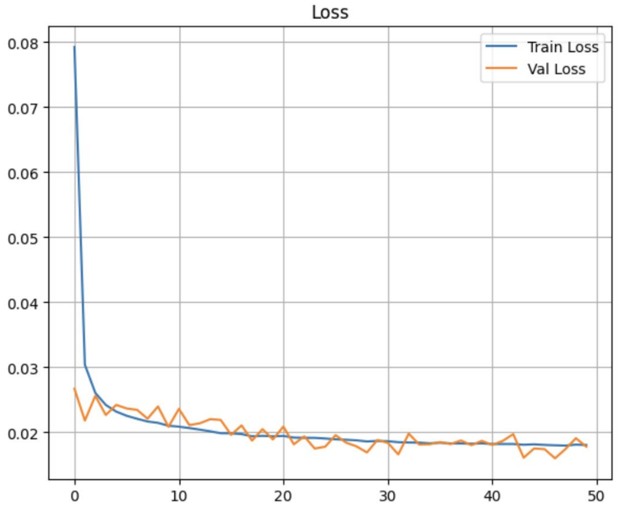
Model	MAE	MSE	RMSE	R^2
RF (Baseline)	0.0734	0.0248	0.1576	0.9906
RF (Tuned)	0.0744	0.0237	0.1541	0.9910

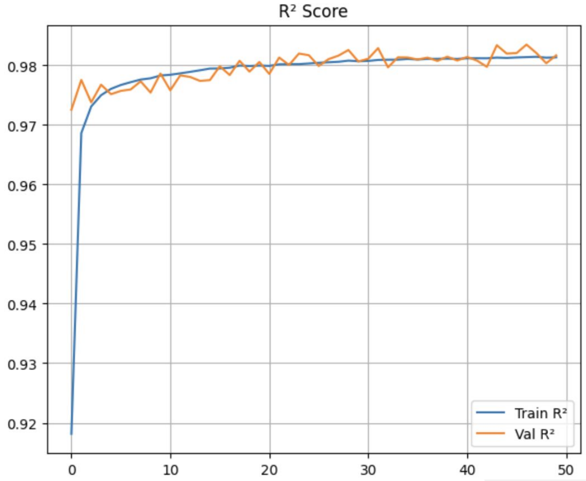
FNN 1 - Simple Model

Layer	Units / Details
Input Layer	Input shape: (input_shape,)
Dense Layer 1	128 units, ReLU activation
Dropout 1	0.3
Dense Layer 2	64 units, ReLU activation
Dropout 2	0.2
Dense Layer 3	32 units, ReLU activation
Output Layer	1 unit (no activation, regression)
Optimizer	Adam (learning rate = 0.001)
Epochs	50

FNN 1 Results

MAE	MSE	RMSE	R2
0.145	0.047	0.216	0.982



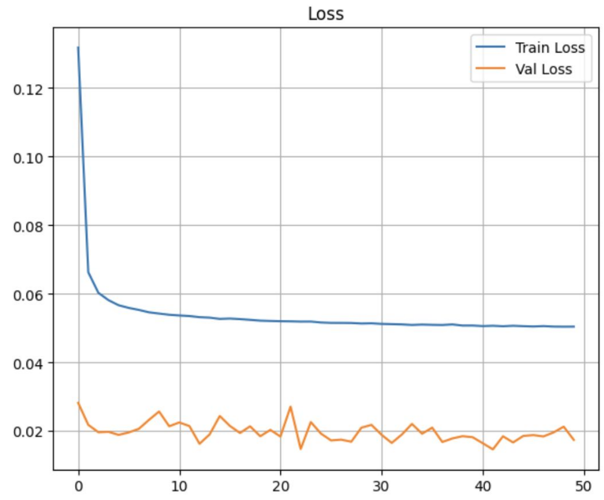


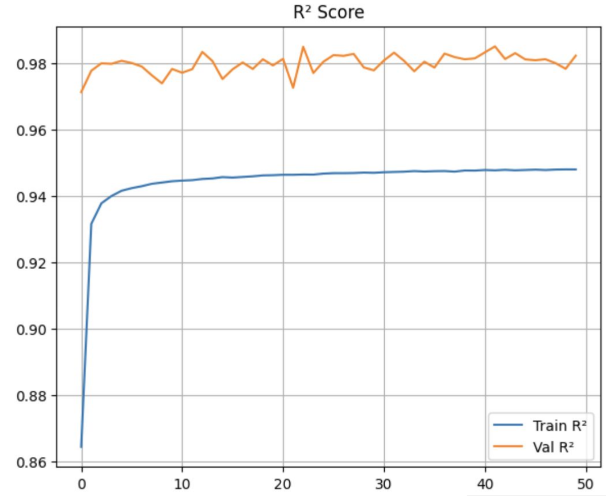
FNN 2 - Complex Model

Layer	Units / Details
Dense Layer 1	512 units, ReLU activation
Batch Normalization 1	-
Dropout 1	0.3
Dense Layer 2	256 units, ReLU activation
Batch Normalization 2	-
Dropout 2	0.3
Dense Layer 3	128 units, ReLU activation
Batch Normalization 3	-
Dropout 3	0.3
Dense Layer 4	64 units, ReLU activation
Batch Normalization 4	-
Dropout 4	0.3
Dense Layer 5	32 units, ReLU activation
Batch Normalization 5	-
Dropout 5	0.3
Output Layer	1 unit (no activation, regression)
Optimizer	Adam (learning rate = 0.001)
Epochs	50

FNN 2 Results

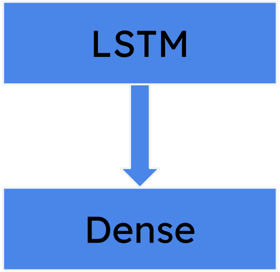
MAE	MSE	RMSE	R2
0.131	0.045	0.215	0.983





LSTM - Model

LSTM model Architecture



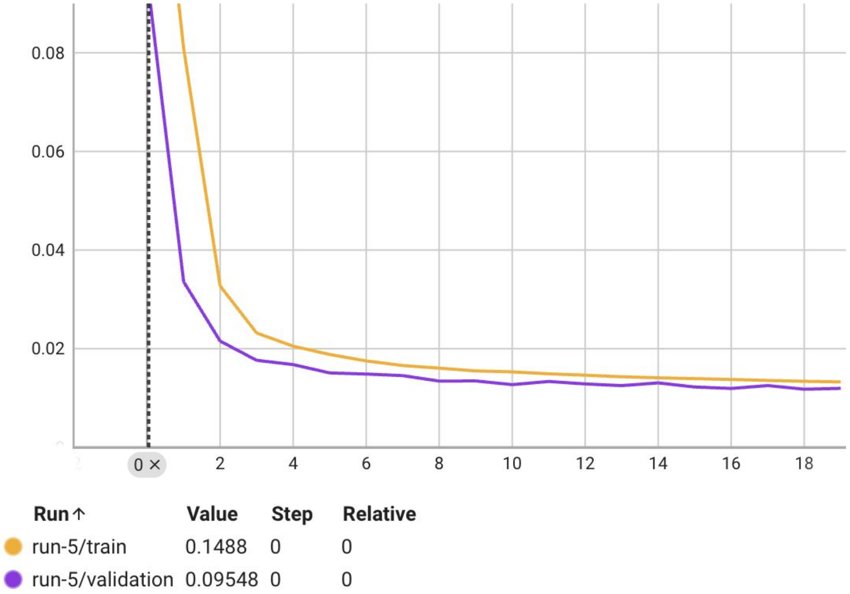
Hyperparameter	Possible Values
Num Units	32, 64, 128
Recurrent Dropout	0.0, 0.2

Units	Rec. Dropout	MAE	MSE	RMSE	R^2
32	0.0	0.111	0.041	0.203	0.984
32	0.2	0.105	0.039	0.198	0.985
64	0.0	0.104	0.037	0.191	0.986
64	0.2	0.098	0.034	0.184	0.987
128	0.0	0.096	0.032	0.179	0.988
128	0.2	0.093	0.032	0.180	0.988

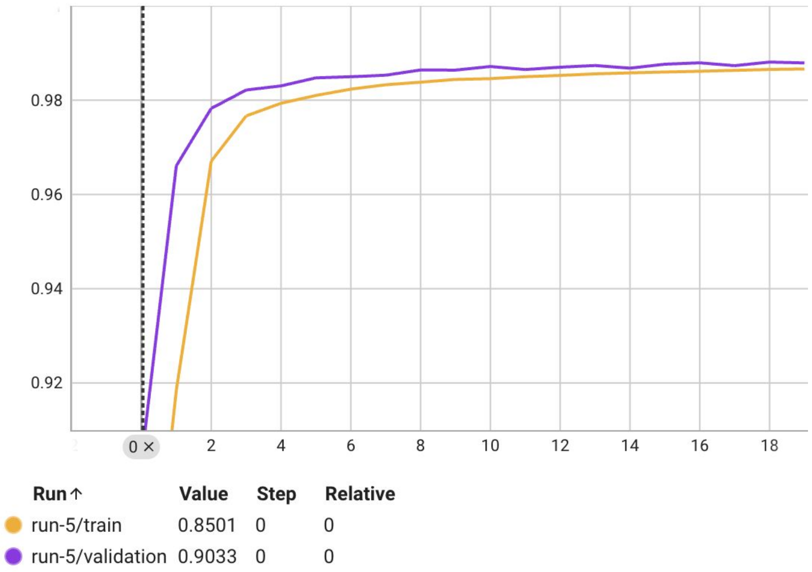
LSTM - Best Model

Units	Rec. Dropout	MAE	MSE	RMSE	R^2
128	0.2	0.093	0.032	0.180	0.988

Loss function over epoch



R^2 function over epoch



Model Comparison

Model	MAE	MSE	RMSE	R^2
Random Forest (Tuned)	0.0744	0.0237	0.1541	0.9910
FNN 1	0.1445	0.0467	0.2161	0.9823
FNN 2	0.1312	0.0452	0.2125	0.9829
FNN with Grouped Features	0.1251	0.0430	0.2073	0.9837
LSTM (Best)	0.0930	0.0320	0.1800	0.9880

Conclusion

The goal of this project was to develop a model capable of predicting energy consumption for the next time step based on building characteristics, weather conditions, and previous lags, across different types of buildings.

The work involved significant effort in data preparation and cleaning, which was crucial for achieving the desired accuracy in the results and ensuring that the model could generalize well across different building conditions and types.

Ultimately, among all the models tested, the LSTM (Long Short-Term Memory) model achieved the best results, as expected, due to its ability to capture temporal and sequential dependencies in the data, which is essential for effectively predicting energy consumption.

Thanks!

Do you have any questions?

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**