Data Analysis of Cooling and Heating Loads of Buildings

By: Anthony

Dataset: "Energy Efficiency" dataset from *UCI Machine Learning Repository*

Agenda

1. Problem Statement
2. Stakeholders
3. Value Propositions
4. Machine Learning
   I. Data Visualization
   II. Feature Engineering
   III. Model Selection & Training
   IV. Model Evaluation
5. Business Recommendations
6. Limitations & Assumptions
7. Areas for Improvement

1. Problem Statement

- Current Issues
  - **Energy efficiency and costs** have been a matter of concern as the global population becomes increasingly developed.
  - Poor management of energy consumption has **negative impact on the environment**

1. Problem Statement

- To predict the heating and cooling load of buildings based on physical and operational parameters, to aid in **accurate** & **energy efficient designs,** thereby **reducing costs** and **limiting environmental impact**.

2. Stakeholders

1. Building Owners & Energy managers

- Operating the facilities
- Pay for energy usage

2. Architects & Mechanical Engineers

- Design building services and heating/cooling capacity

3. Building Occupants

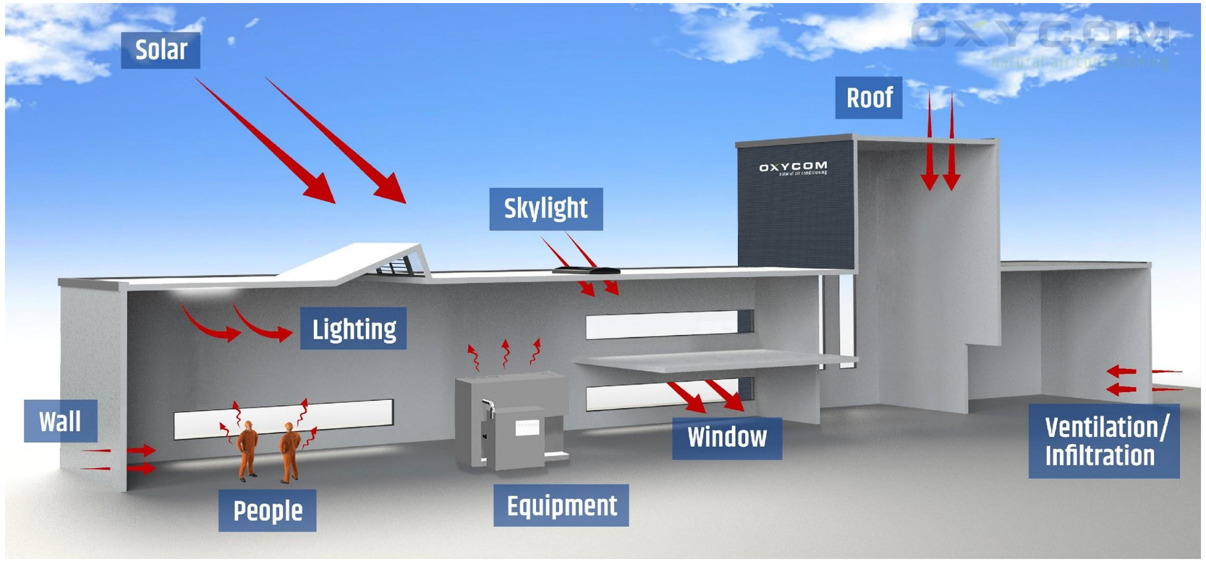
- Comfort

3. Value Proposition

1. Reduced Operating costs
2. Reduced Start-Up costs
3. Improved occupant comfort
4. Low environmental impact

Machine Learning

- Pre-AI



COOLING LOAD CALCULATION SHEET

Project Name : .... Prepare by : .... Revision : ....

Location : .... Approved by : .... Sheet No. : ....

	SPACE USED FOR	Switchgear/CCR Building				
	SIZE WxLxH (M)	40	14	3	LOCATION	
3 & 4	AREA (m2)	560	VOLUME (M3)	1680	MAINDECK	
	ITEM	AREA OR QUANTITY		SUNGAIN OR TEMP. DIFF		FACTOR U VALUE
	SOLAR GAIN-GLASS (Watts)					-
	GLASS	0	m^2 X	341.14	X	0.94
	SKYLIGH	0	m^2 X	341.14	X	0.94
5	SOLAR & TRANS. GAIN-WALLS & ROOF (Watts)					-
	WALL NW	0	m^2 X	19	X	0.34
	WALL SE	0	m^2 X	15	X	0.34
	WALL NE	0	m^2 X	16	X	0.34
	WALL SW	0	m^2 X	17	X	0.34
	ROOF-SUN	0	m^2 X	28	X	0.49
	ROOF-SHADED	0.00	m^2 X	15	X	0.34
	TRANS. GAIN-EXCEPT WALLS & ROOF (Watts)					-
	ALL GLASS	0	m^2 X	12	X	1.07
	PARTITION	0	m^2 X	12	X	0.42
	CEILING	0	m^2 X	12	X	0.42
	FLOOR	0.00	m^2 X	12	X	0.42
6	INFILTRATION	0	L/s X	12	X	1.2
3 & 7	INTERNAL HEAT (Watts)					-
	PEOPLE	0	people X	72	W	0
	EQUIP HEAT	0.00	kW X	1000	W/m2	0
	LIGHT	0.00	m2 X	12.19	W/m2	0
	APPLIANCE ETC.					
	ADDITIONAL HEAT GAIN					
2 & 3	SUB TOTAL 1 (Watts)					-
	STORAGE	0	m2 X	3	X	0.4
	SUB TOTAL 2 (Watts)					-
7	SAFETY FACTOR	0		%		
	ROOM SENSIBLE HEAT (Watts)					186,174
	SUPPLY DUCT HEAT GAIN %	+ 3	SUPPLY DUCT LEAK LOSS %	+ 2	FAN HG	5 %
8	OUTDOOR AIR	1153.0	L/s X	12	oC X 0.06	BF X 1.2
	EFFECTIVE ROOM SENSIBLE HEAT(Watt)					205,788
6	LATENT HEAT (Watt)					360
	INFILTRATION	0	L/s	15	g/kg	
3 & 7	PEOPLE	6	PEOPLE X	60	W	
	STEAM	0	LB/HR X 1050			
	APPLIANCE ETC.		0		X	0
	ADDITIONAL HEAT GAIN		0			
5	VAPOR TRANS.					
7	SAFETY FACTOR	5	%			
	ROOM LATENT HEAT (Watt)					378
	SUPPLY AIR DUCT LEAK LOSS			2	%	
8	OUTDOOR AIR	1153	L/s X	23.9	g/kg X 0.06	BF X 3.0
	EFFECTIVE ROOM LATENT HEAT (Watt)					5,346
	EFFECTIVE ROOM TOTAL HEAT (Watt)					211,134
	OUTDOOR AIR HEAT					93,317
	SENSIBLE :	1153.0	L/s X	12	oC	X (1- 0.06 BF) X 1.2
	LATENT :	1153	L/s X	23.9	g/kg	X (1- 0.06 BF) X 3.0
7	RETURN DUCT HEAT GAIN %	5		RETURN DUCT + LEAK GAIN %		2
	ROOM SENSIBLE HEAT FACTOR			RSHF		1.00
	GRAND TOTAL HEAT FACTOR			GTHF		0.71
	SUB TOTAL 4 (Watt)					100,780
	GRAND TOTAL HEAT (Watt)					311,914

	ESTIMATE FOR	15.00 PM	LOCAL TIME SUN TIME	PEAK LOAD	12.00 AM	LOCAL TIME SUN TIME
	HOUR OF OPERATION			24		
2	CONDITIONS	DB	WB	% RH	DP	g/kg
	OUTDOOR (OA)	35	33.4	90	-	32.6
	ROOM (RM)	23	16.2	50	-	8.7
	DIFFERENCE	12	-	-	-	23.9

6 — VENTILATION — OUTDOOR AIR

6 PEOPLE X 12 L/s per pax = 72

560.00 m2 X 0.3 L/s per m2 = 168

L/s VENTILATION = 168

EXFILTRATION — ROOM PRESSURIZATION 50 Pa

=

Ventilation Required =

As per Design Data Input 1210 L/s

OUTDOOR AIR THRU APPARATUS 1153.0 L/s_{OA}

8 — APPARATUS DEWPOINT

ESHF — EFFECTIVE SENS. HEAT = $\frac{\text{EFFECTIVE ROOM SENS. HEAT}}{\text{EFFECTIVE ROOM TOTAL HEAT}}$ = 0.975

ADP — INDICATED ADP 11.8 oC , SELECTED AD 11.8 oC

DEHUMIDFIED AIR QUANTITY

TEMP. RISE — (1 - 0.06 BF) X (T_{RM} 23 oC - T_{ADP} 11.8 oC) = 10.528 K

DEHUM FLOW — $\frac{\text{EFFECTIVE ROOM SENS. HEAT}}{1.2 \text{ X } 10.528\ K_{TEMP.RISE}}$ = 16,332 L/s_{DA}

OUTLET TEMP. — $\frac{\text{ROOM SENS. HEAT}}{1.2 \text{ X } 16332\ L/s_{DA}}$ = 9.50 $K_{(RM-OUTLET AIR)}$

SUPPLY AIR QUANTITY

SUPPLY AIR FLOW — $\frac{\text{ROOM SENS. HEAT}}{1.2 \text{ X } 9.50\ k_{DESIRE DIFF}}$ = 16,332 L/s_{SA}

BYPASS AIR FLOW — 16332 L/s_{SA} - 16,313 L/s_{DA} = 19 L/s_{BA}

RESULTING ENT & LVG CONDITIONS AT APPARATUS

MIXING AIR — T_{RM} 23 oC + 0.07 $\frac{OA}{SA}$ X (T_{OA} 35 oC - T_{RM} 23 oC) = T_{EDB} 23.8 oC

LEAVING AIR — T_{ADP} 11.8 oC 0.06 BF X (T_{EDB} 24 oC - T_{ADP} 11.8 oC) = T_{LDB} 12.5 oC

FROM PSYCH. CHART : T_{EWB} 18 oC , T_{LWB} 12.2 oC

GRAND TOTAL COOLING LOAD

DIFF. ENTHALPY — ENTER AIR ENTHALPY 50.67 kj/kg; LEAVING AIR ENTHALPY 34.67 kj/kg; DIFFERENT ENTHALPY 16 kj/kg

TOTAL COOLING LOAD — 16,332 L/s x 1.2 x 16 = 313582 Watt

MARGIN FACTOR = 10%

COOLING LOAD = 344940 Watt

* IF THIS DIFF TEMP IS TOO HIGH, DETERMINE SUPPLY CFM FOR DESIRED DIFFERENCE BY SUPPLY AIR QUANTITY FORMLA.

*WHEN BYPASSING MIXTURE OF OUTDOOR AND RETURN AIR, USE SUPPLY CFM WHEN BYPASSING RETURN AIR ONLY, USE DEHUMIFIED CFM.

4. Machine Learning – I. Data Visualization

- Features

Name	Description
X1	Relative Compactness
X2	Surface Area
X3	Wall Area
X4	Roof Area
X5	Overall Height
X6	Orientation
X7	Glazing Area
X8	Glazing Area Distribution

- Targets

Name	Description
Y1	Heating Load
Y2	Cooling Load

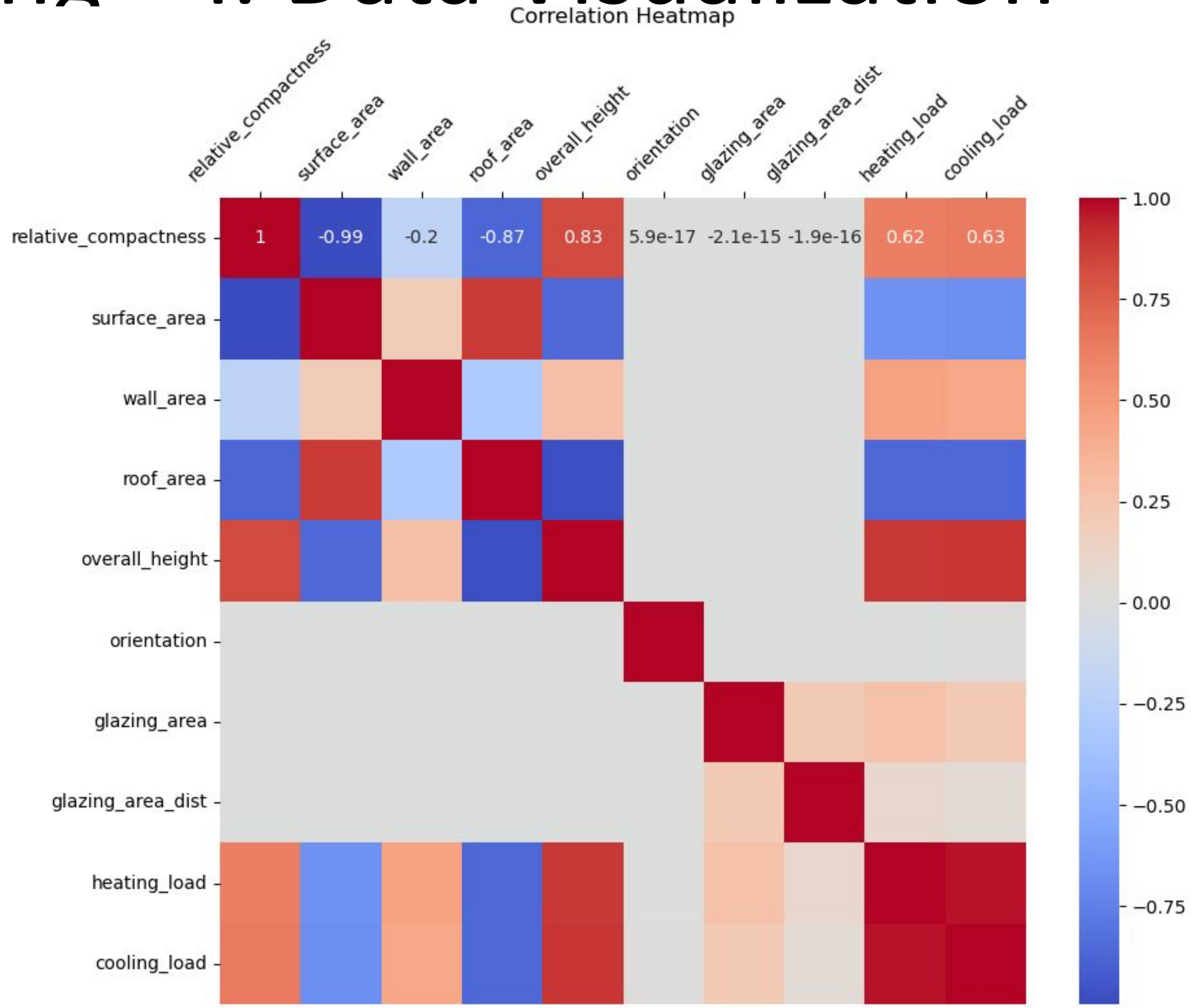
4. Machine Learning – I. Data Visualization

- Dataset
  - 768 points

```
RangeIndex: 768 entries, 0 to 767
Data columns (total 10 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   X1      768 non-null    float64   Continuous
 1   X2      768 non-null    float64   Continuous
 2   X3      768 non-null    float64   Continuous
 3   X4      768 non-null    float64   Continuous
 4   X5      768 non-null    float64   Continuous
 5   X6      768 non-null    int64        Integer
 6   X7      768 non-null    float64   Continuous
 7   X8      768 non-null    int64        Integer
 8   Y1      768 non-null    float64   Continuous
 9   Y2      768 non-null    float64   Continuous
```

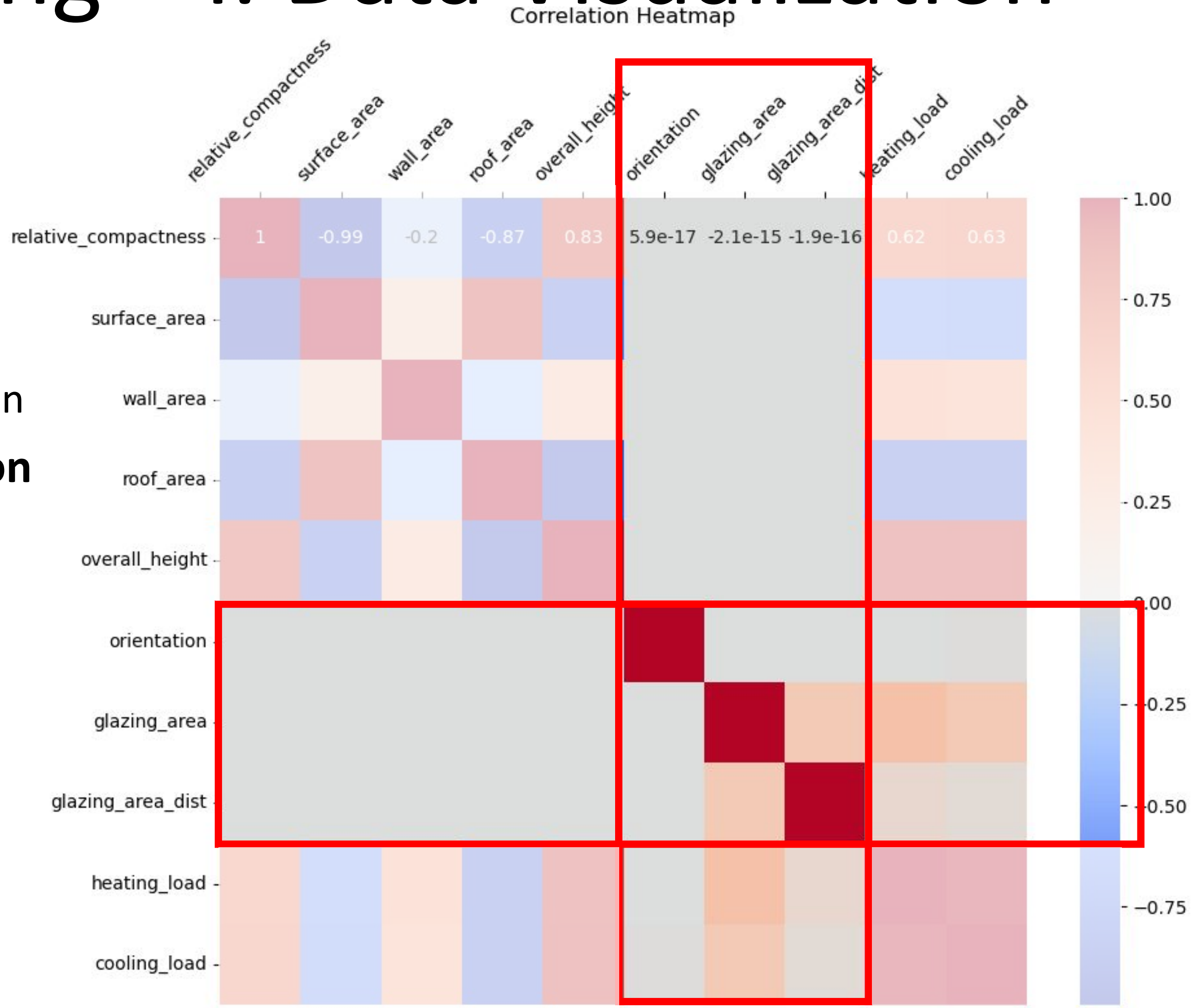
4. Machine Learning – I. Data Visualization

- Correlation Heatmap



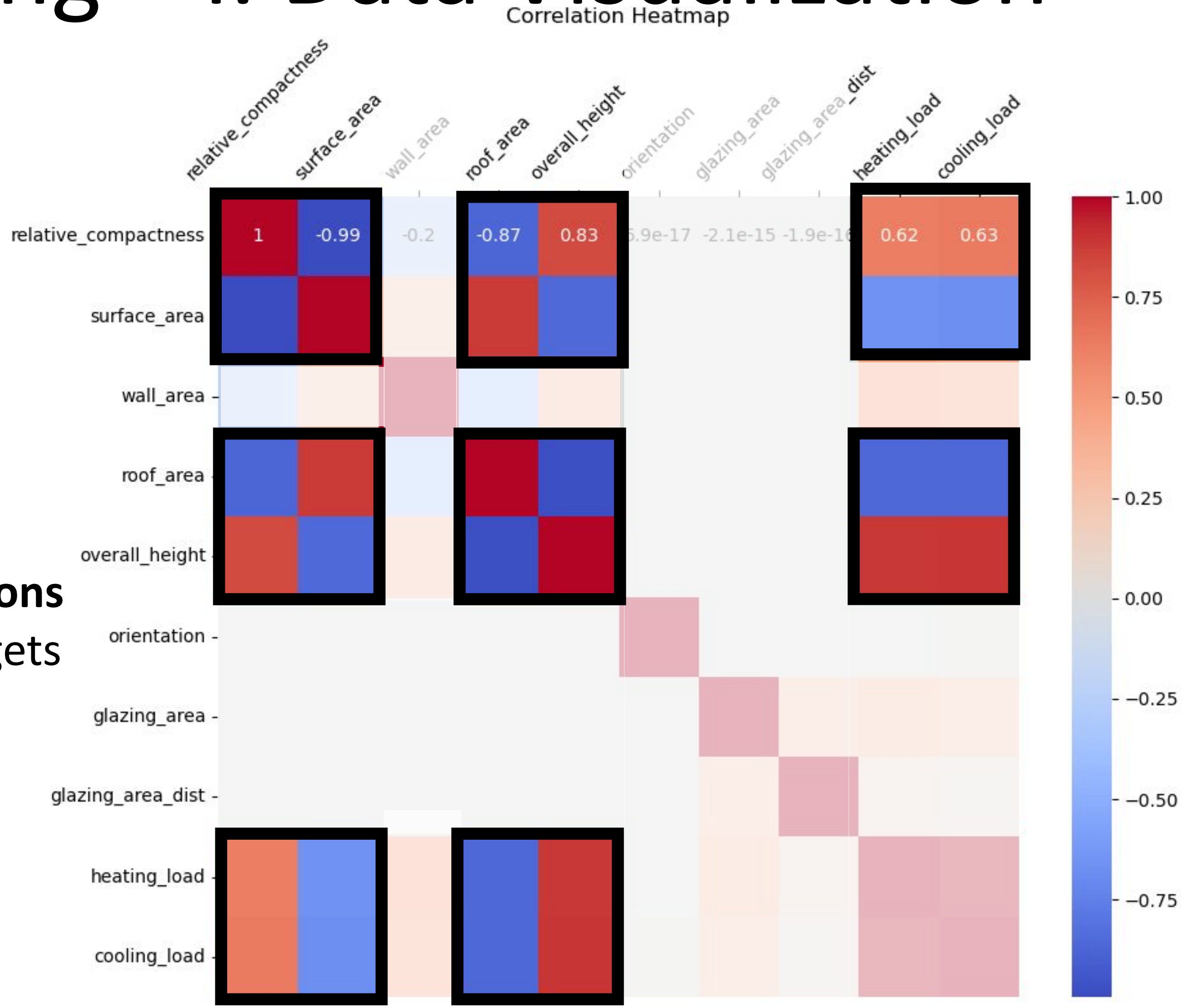
4. Machine Learning – I. Data Visualization

- Correlation Heatmap
  - Key Observations
    - X7 Glazing Area
    - X8 Glazing Area Distribution
  - Features have **low correlation** to other features and the targets



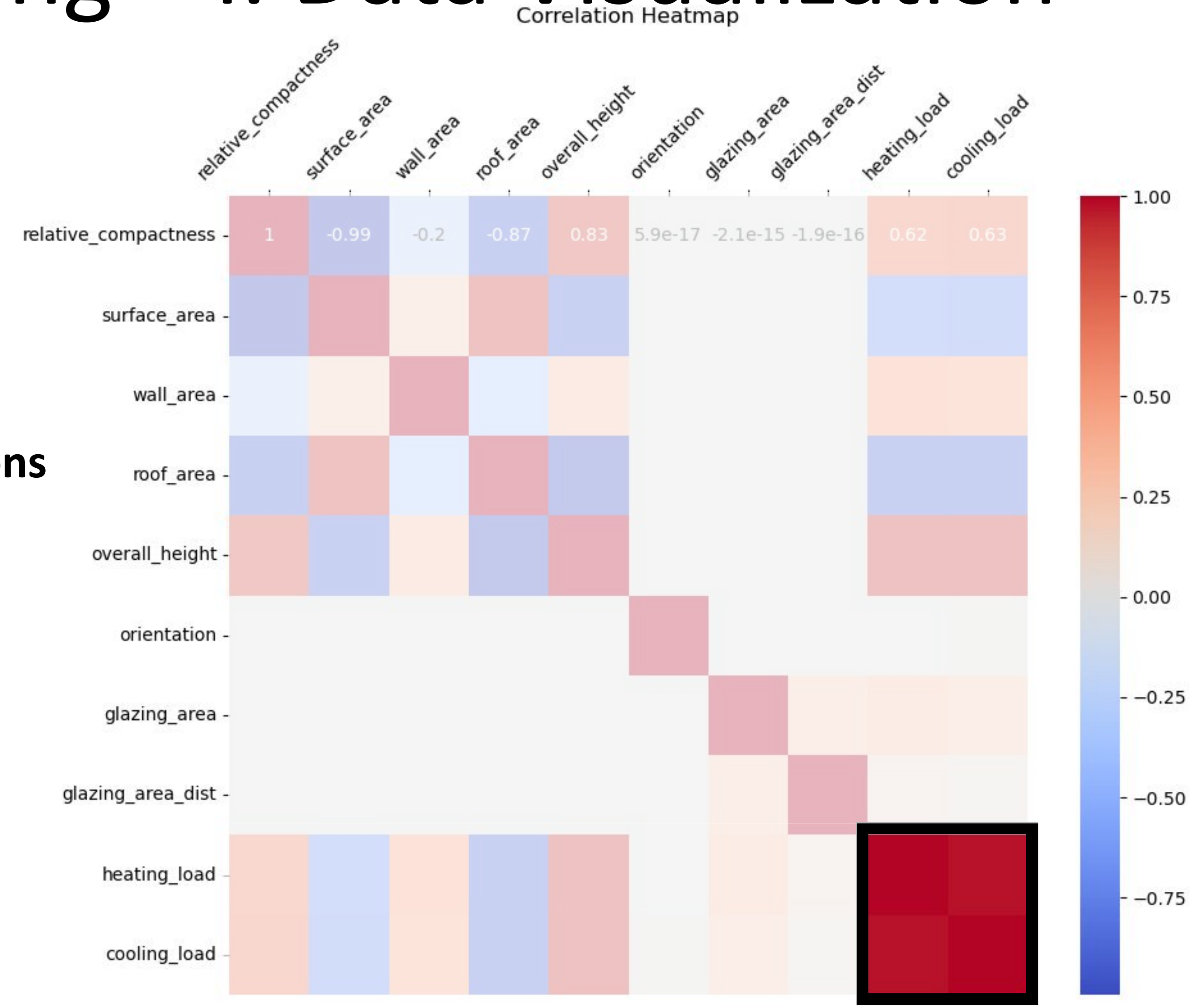
4. Machine Learning – I. Data Visualization

- Correlation Heatmap
  - Key Observations
    - X1 Relative Compactness
    - X2 Surface Area
    - X4 Roof Area
    - X5 Overall Height
  - Features have **high correlations** with each other and the targets



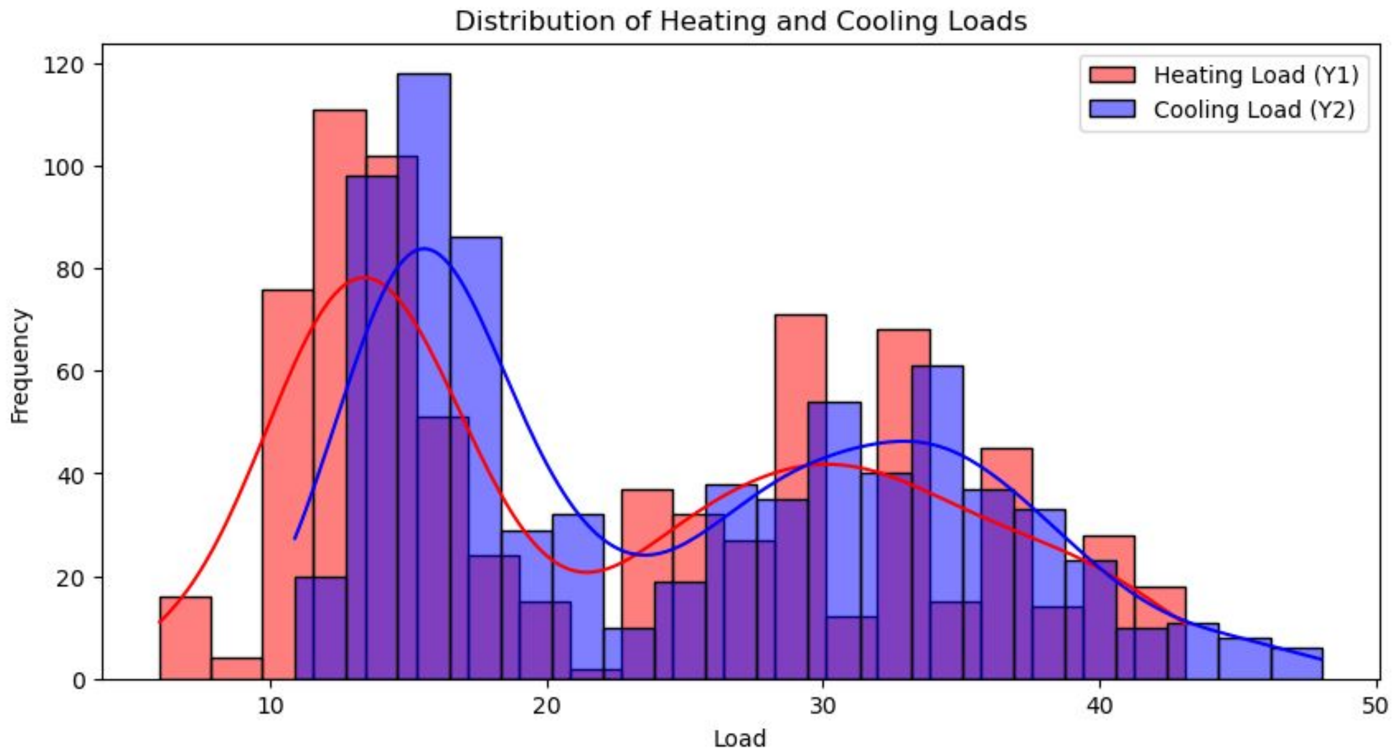
4. Machine Learning – I. Data Visualization

- Correlation Heatmap
  - Key Observations
    - Y1 Heating Load
    - Y2 Cooling Load
  - Targets have **high correlations** with each other



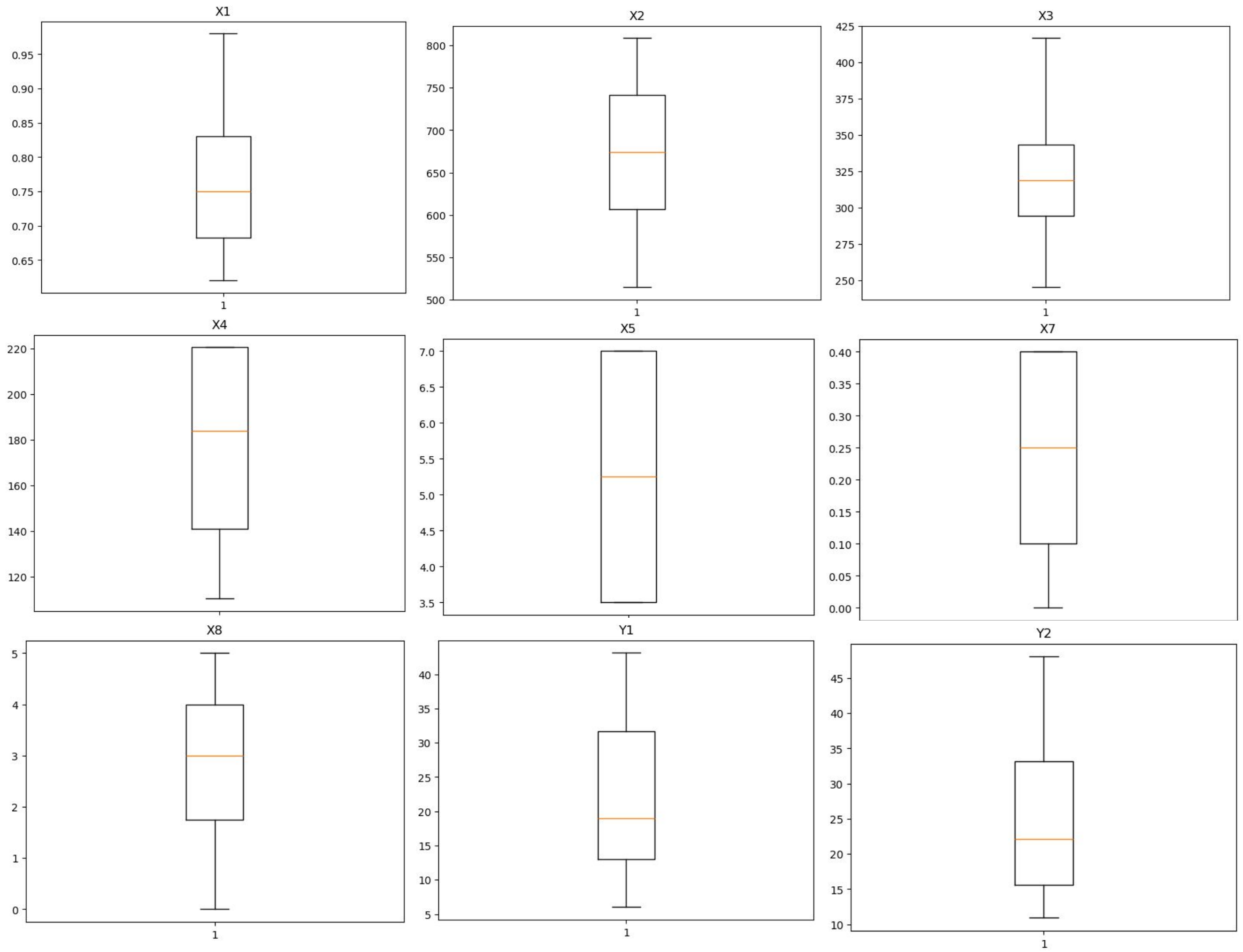
4. Machine Learning – I. Data Visualization

- Histplot of Targets
  - Cooling load is **higher** than Heat Load, but have similar distributions.



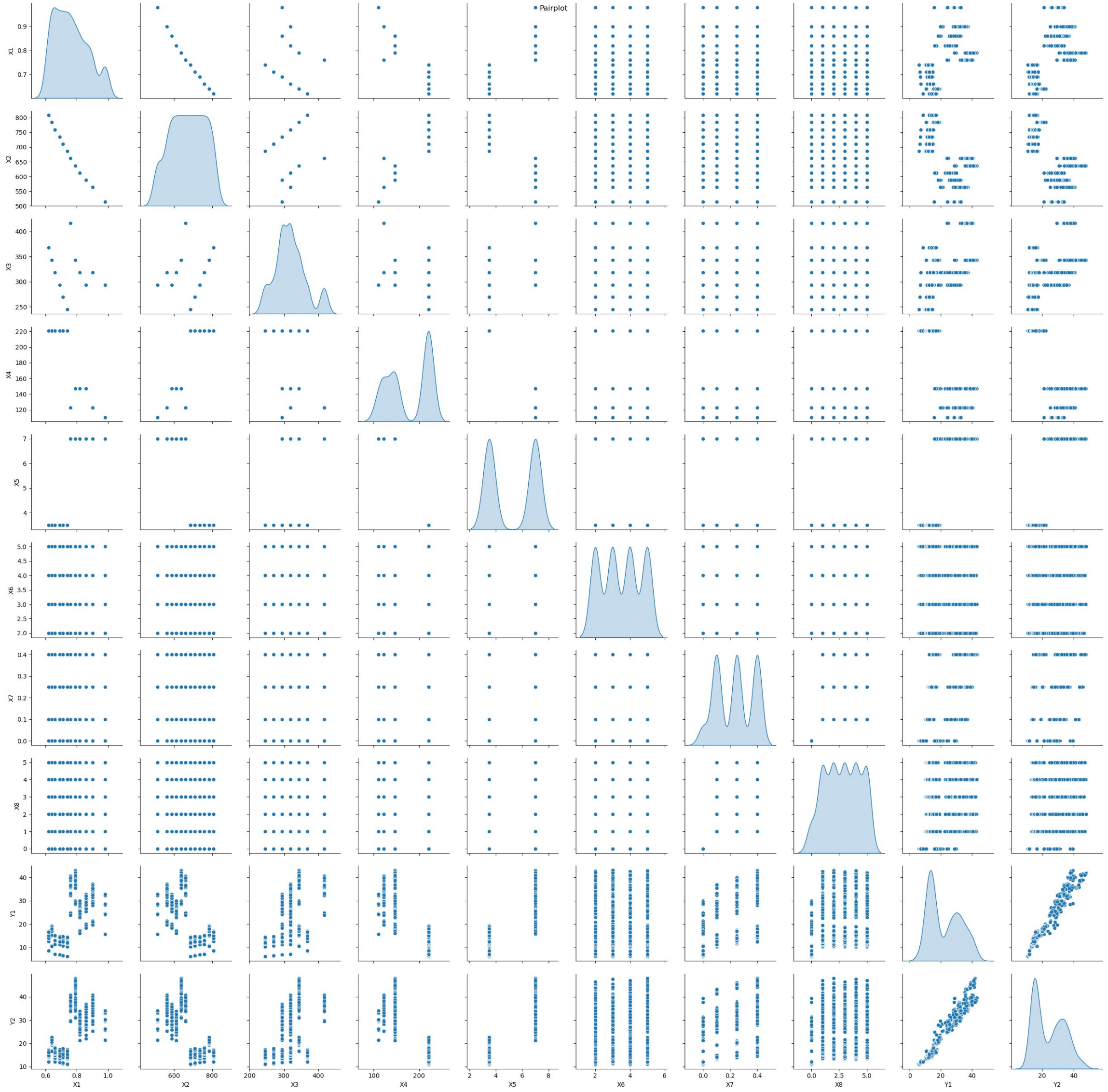
4. Machine Learning – I. Data Visualization

- Box Plot
  - No outliers



4. Machine Learning – I. Data Visualization

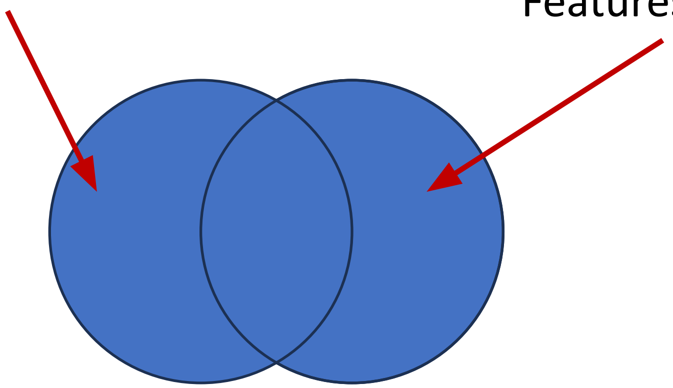
- Pair Plot
  - **Non-linear** relationships between features and targets



4. Machine Learning – II. Feature Engineering

- Feature Selection
  - **Lasso Regressor** for each output
  - Select most important Features

Selected Features = **Union of** individual output **selected features**



4. Machine Learning – II. Feature Engineering

- Lasso Selected Features

Name	Description
~~X1~~	~~Relative Compactness~~
X2	Surface Area
X3	Wall Area
X4	Roof Area
X5	Overall Height
X6	Orientation
X7	Glazing Area
X8	Glazing Area Distribution

- Targets

Name	Description
Y1	Heating Load
Y2	Cooling Load

4. Machine Learning – III. Model Selection

- Initial Selection:
  - **MultiOutput**Regressor() Wrapped over
    - LinearRegression()
    - DecisionTreeRegressor()
    - RandomForestRegressor()
    - KNeighborsRegressor()

- To be cross validated
  - Obtain Model selection with lowest Mean Absolute Error

4. Machine Learning – III. Model Selection

- Selected Model (Lowest MAE):
  - **MultiOutput**Regressor()
    - LinearRegression() 2.39
    - DecisionTreeRegressor() 0.75
    - RandomForestRegressor() 0.67
    - KNeighborsRegressor() 1.63

4. Machine Learning – III. Model Selection

1. **Split Data** into training & testing sets
2. **Fit** training data into model
3. **Hyperparameter Tuning** using GridSearch
4. **Evaluation**

4. Machine Learning – IV. Model Evaluation

<table>
<tr>
<td>Mean Squared Error

(Lower is better)</td>
<td>

	MSE
Y1	**0.2870**
Y2	**2.7245**

</td>
<td>- High degree of accuracy for Heat Load predictions

- Good degree of accuracy for Cooling Load predictions</td>
</tr>
<tr>
<td>R^2 Score

(Higher is better)</td>
<td>

	R^2
Y1	**0.9972**
Y2	**0.9706**

</td>
<td>- Fits the data very well

- High Predictive power</td>
</tr>
</table>

4. Machine Learning – IV. Model Evaluation

Mean Absolute Percentage Error

(*Lower is better*)

	MAE	MAPE
Y1	0.4017	**1.81%**
Y2	1.0573	**3.55%**

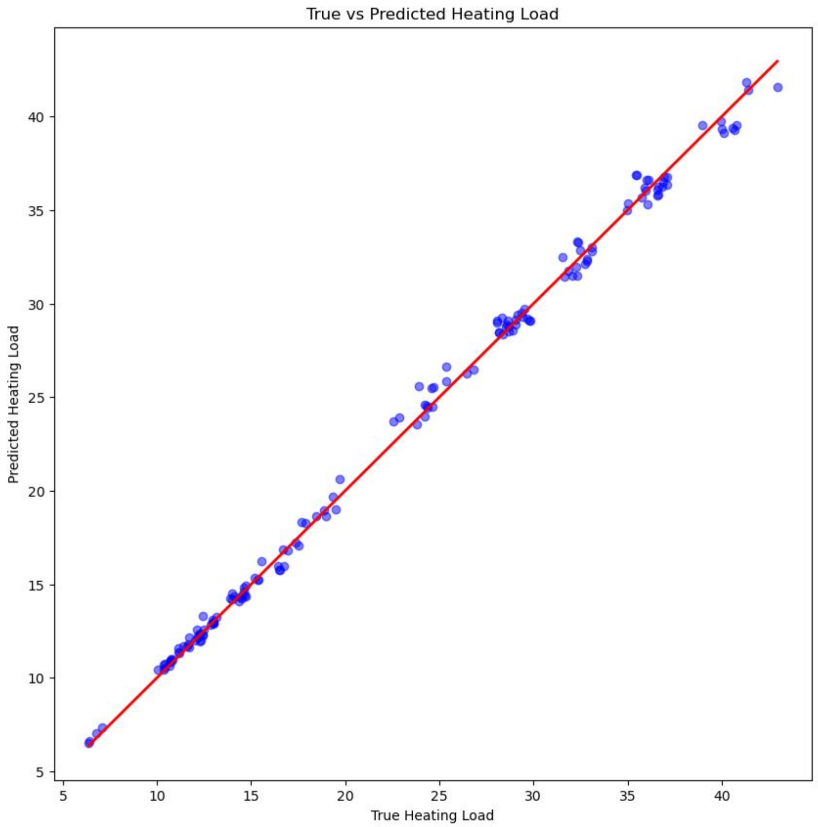
- **Very Low percentage error** for Heat Load predictions
- **Low percentage error** for Cooling Load predictions

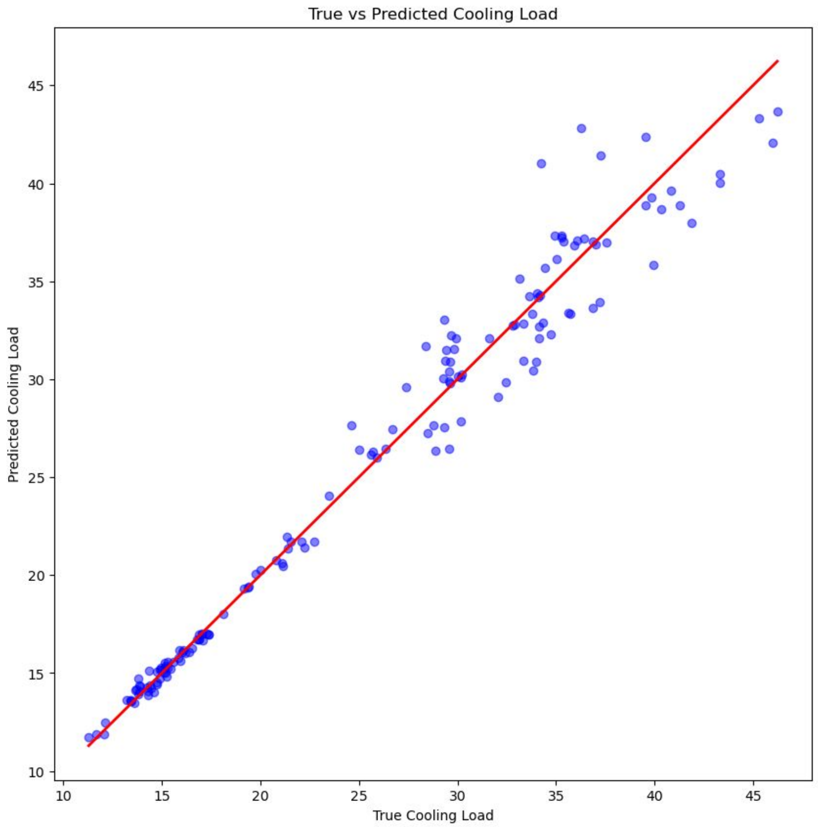
4. Machine Learning – IV. Model Evaluation

- Predicted data vs True data

Heating Load

Cooling Load





4. Machine Learning – IV. Model Evaluation

Conclusion

∴ Model is **reliable** and **accurate** in predictions of Heating and Cooling Load.

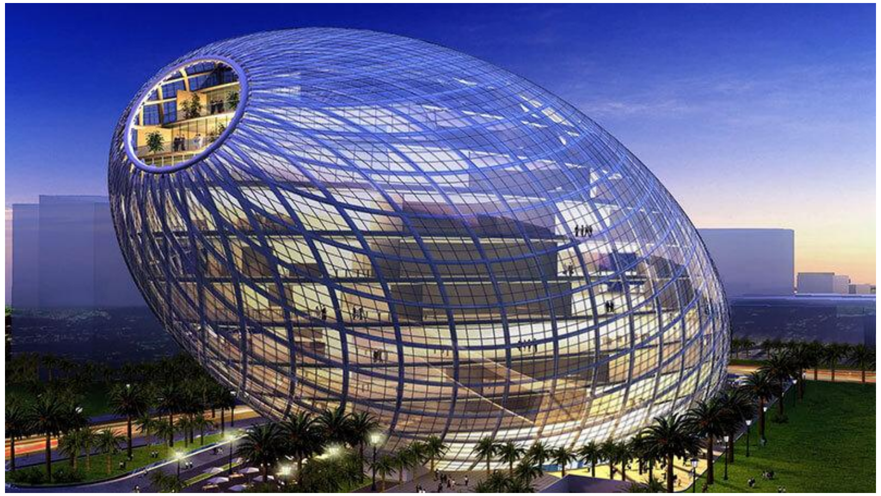
5. Business Recommendations

- Building Owners
  - Upgrades & Retrofits
  - Cost Reduction Strategies
  - Achieve Green Mark Energy Efficiency

- Architects & Engineers
  - Optimized Building Designs
  - Validation of engineering calculations
  - Improved budgeting
  - Reduced manhours & cost

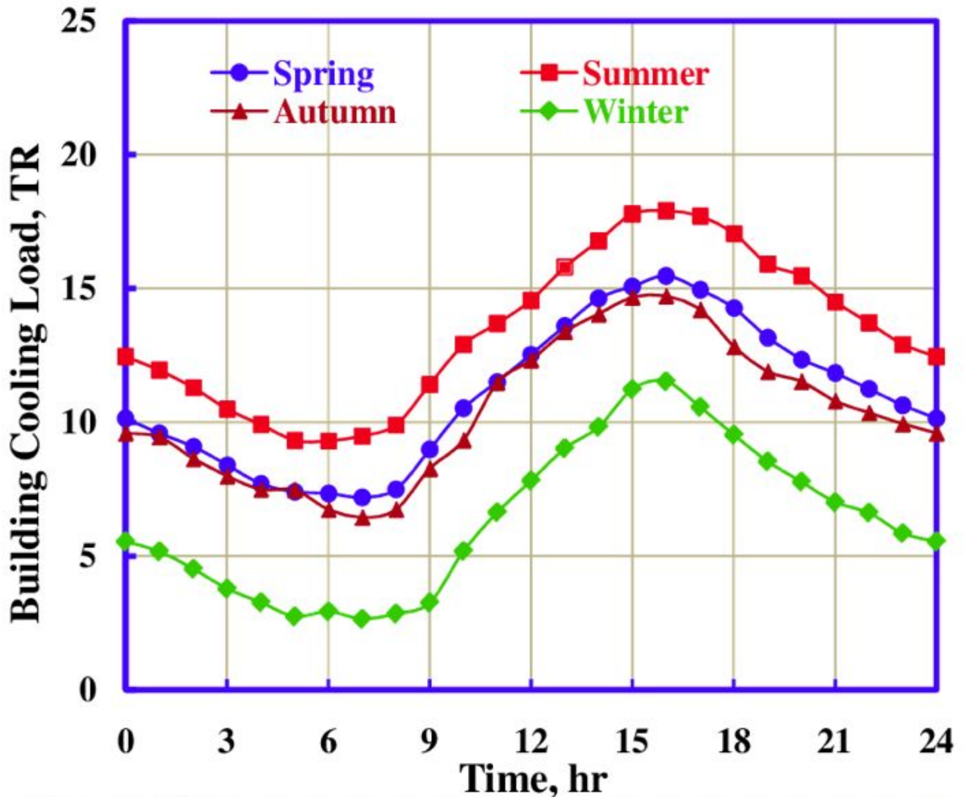
6. Limitations & Assumptions

- Limitations
  - **Limited Dataset** – Data represents a sample of all buildings in the world
    - (Architects love designing ***odd-shaped*** buildings)
  - **Limited features** – e.g. Location, climate, types of building



6. Limitations & Assumptions

- Assumptions
  - **Time variable**
    - Day time vs Night time
    - Seasons, climate



7. Areas for Improvement

- Additional features
  - Climate and weather data
- Larger dataset
- Incorporate engineering calculations

Thank you for your attention!

Any questions?