National College of Ireland

Configuration Manual

MSc Research Project
Master's in Data Analytics

Bharath Raparthi
23417226

School of Computing
National College of Ireland

Supervisor: Mr. Eamon Nolan

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Bharath Raparthi
Student ID:	23417226
Programme:	Master's in Data Analytics **Year:** 2025-26
Module:	Research Practicum
Lecturer:	Eamon Nolan
Submission Due Date:	11/12/25
Project Title:	Multi-Type Building Energy Consumption Prediction Using Machine Learning and IoT Sensor Data

Word Count: 144 **Page Count:** 4

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Bharath Raparthi

Date: 11/12/25

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

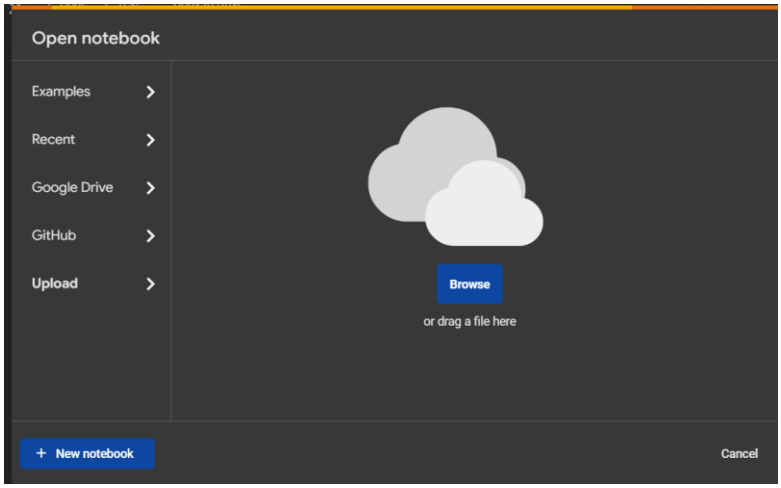
Bharath Raparthi
23417226

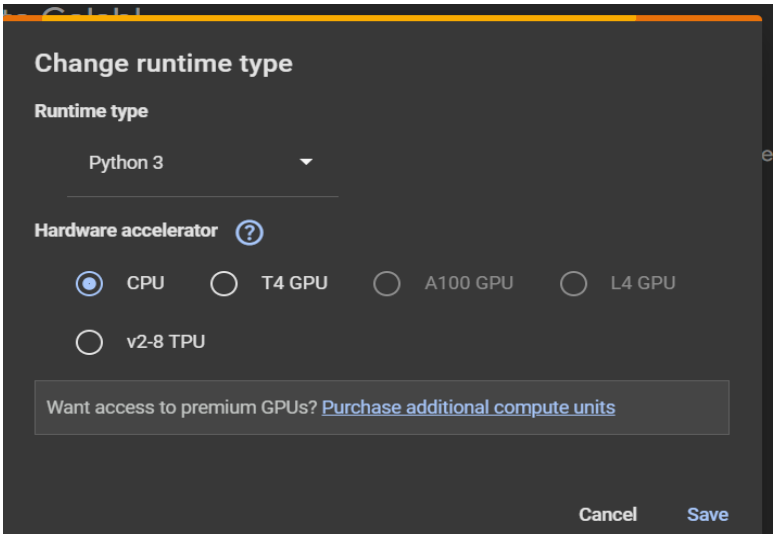
1 Linking Google Colab with google drive

Step 1: Open google and then type google Coloab
Step 2: Now once open Sign into the google account
Step 3: Now upload the file

1. Upload the modeling file

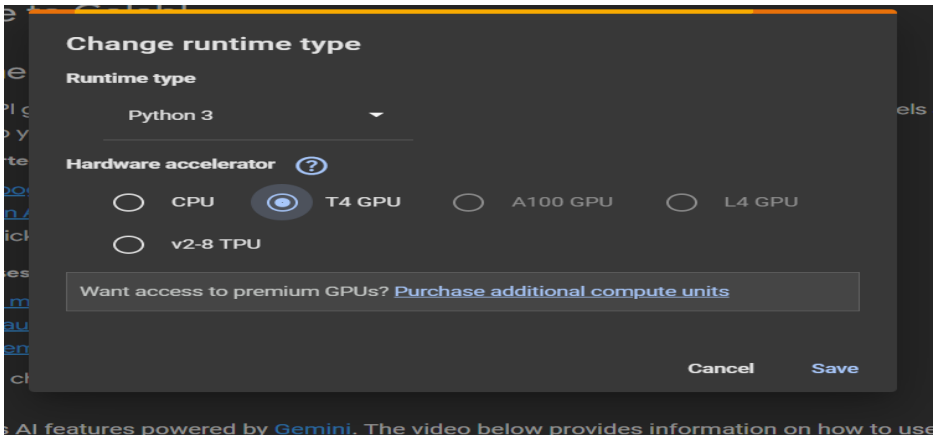
Step 4 : Now enable the GPU acceleration
Step 5: Change the Runtime





2 Setting up the Session

Step 1: Now open the google colab session
Step 2: Now sign into your account
Step 3: Start the New Notebook
Step 4: Now Activate the GPU
Step 5: Upload the Modelling file



3 Import of the data

Step 1: Import the train and test data from drive
Step 2: Import the configurations

```
train_df=pd.read_csv("/content/drive/MyDrive/Bharath_Energey/final_train_dataset.csv")
```

```
test_df=pd.read_csv("/content/drive/MyDrive/Bharath_Energey/final_test_dataset.csv")
```

```
train_df.columns
```

```
Index(['building_id', 'meter', 'timestamp', 'meter_reading', 'site_id',
       'primary_use', 'square_feet', 'year_built', 'floor_count',
       'air_temperature', 'cloud_coverage', 'dew_temperature',
       'precip_depth_1_hr', 'sea_level_pressure', 'wind_direction',
       'wind_speed'],
      dtype='object')
```

```
np.random.seed(RANDOM_SEED)

# Configuration
CONFIG = {
    'train_path': '/content/drive/MyDrive/Bharath_Energey/final_train_dataset.csv',
    'test_path': '/content/drive/MyDrive/Bharath_Energey/final_test_dataset.csv',
    'output_path': '/content/drive/MyDrive/Bharath_Energey/',
    'validation_split_date': '2016-11-01',  # Last 2 months for validation
    'meter_types': {0: 'electricity', 1: 'chilledwater', 2: 'steam', 3: 'hotwater'},
    'quantiles': [0.1, 0.5, 0.9],  # For uncertainty quantification
}

print("Configuration loaded successfully")

Configuration loaded successfully
```

4 Results

```
ELECTRICITY:
--------------------------------------------------
XGBoost:    RMSE=  200.11 | MAE=   30.98 | MAPE=  60.9%
LightGBM:   RMSE=  172.17 | MAE=   32.63 | MAPE=  55.7%
Ensemble:   RMSE=  189.32 | MAE=   31.09 | MAPE=  57.8% | Coverage=74.5%
Improvement: +5.4% vs XGBoost, -10.0% vs LightGBM

CHILLEDWATER:
--------------------------------------------------
XGBoost:    RMSE= 2551.07 | MAE=  162.14 | MAPE= 561.5%
LightGBM:   RMSE= 2554.61 | MAE=  175.40 | MAPE= 357.1%
Ensemble:   RMSE= 2551.62 | MAE=  164.43 | MAPE= 434.5% | Coverage=75.2%
Improvement: -0.0% vs XGBoost, +0.1% vs LightGBM

STEAM:
--------------------------------------------------
XGBoost:    RMSE=161363.55 | MAE= 2199.13 | MAPE= 875.4%
LightGBM:   RMSE=161352.73 | MAE= 2184.31 | MAPE= 721.5%
Ensemble:   RMSE=161358.32 | MAE= 2172.46 | MAPE= 777.2% | Coverage=78.5%
Improvement: +0.0% vs XGBoost, -0.0% vs LightGBM

HOTWATER:
--------------------------------------------------
XGBoost:    RMSE= 3330.06 | MAE=  397.88 | MAPE=1050.4%
LightGBM:   RMSE= 3304.73 | MAE=  396.19 | MAPE= 841.2%
Ensemble:   RMSE= 3325.04 | MAE=  396.59 | MAPE= 919.4% | Coverage=64.6%
Improvement: +0.2% vs XGBoost, -0.6% vs LightGBM
```

```
================================================================================
COMPREHENSIVE MODEL COMPARISON SUMMARY
================================================================================

Model Performance by Meter Type:
--------------------------------------------------------------------------------
electricity     | XGBoost    | RMSE:     200.11 | MAE:      30.98 | MAPE:    60.9%
electricity     | LightGBM   | RMSE:     172.17 | MAE:      32.63 | MAPE:    55.7%
electricity     | Ensemble   | RMSE:     189.32 | MAE:      31.09 | MAPE:    57.8% | Coverage: 74.5%
chilledwater    | XGBoost    | RMSE:    2551.07 | MAE:     162.14 | MAPE:   561.5%
chilledwater    | LightGBM   | RMSE:    2554.61 | MAE:     175.40 | MAPE:   357.1%
chilledwater    | Ensemble   | RMSE:    2551.62 | MAE:     164.43 | MAPE:   434.5% | Coverage: 75.2%
steam           | XGBoost    | RMSE:  161363.55 | MAE:    2199.13 | MAPE:   875.4%
steam           | LightGBM   | RMSE:  161352.73 | MAE:    2184.31 | MAPE:   721.5%
steam           | Ensemble   | RMSE:  161358.32 | MAE:    2172.46 | MAPE:   777.2% | Coverage: 78.5%
hotwater        | XGBoost    | RMSE:    3330.06 | MAE:     397.88 | MAPE:  1050.4%
hotwater        | LightGBM   | RMSE:    3304.73 | MAE:     396.19 | MAPE:   841.2%
hotwater        | Ensemble   | RMSE:    3325.04 | MAE:     396.59 | MAPE:   919.4% | Coverage: 64.6%

--------------------------------------------------------------------------------
AVERAGE PERFORMANCE ACROSS ALL METER TYPES:
--------------------------------------------------------------------------------
XGBoost Average RMSE:  41861.19
LightGBM Average RMSE: 41846.06
Ensemble Average RMSE: 41856.07

Ensemble improvement: 0.0% better than XGBoost
                      -0.0% better than LightGBM
```

```
Using best models based on validation performance:
  electricity: Using LightGBM (RMSE=172.17)
  chilledwater: Using XGBoost (RMSE=2551.07)
  steam: Using LightGBM (RMSE=161352.73)
  hotwater: Using LightGBM (RMSE=3304.73)

Generating predictions...
  Generated 593,391 predictions for electricity
  Generated 209,306 predictions for chilledwater
  Generated 136,594 predictions for steam
  Generated 60,709 predictions for hotwater

Prediction summary:
  LightGBM predictions: 790,694
  XGBoost predictions: 209,306
  Ensemble predictions: 0

Saved 1,000,000 predictions to predictions.csv

Saving trained models...

================================================
PIPELINE COMPLETE!

Key Deliverables Created:
1. predictions.csv - Final predictions with uncertainty bounds
2. feature_importance.csv - Feature importance rankings
3. results_summary.json - Comprehensive performance metrics
4. feature_importance_by_meter.png - Visualization
5. model_artifacts.pkl - Trained models for deployment
```

5 References

Google Colab. (n.d.). https://colab.research.google.com/

Borrotti, M. (2024). Quantifying Uncertainty with Conformal Prediction for Heating and Cooling Load Forecasting in Building Performance Simulation. *Energies*, *17*(17), 4348. https://doi.org/10.3390/en17174348