

Intermediate Project (Option A):

Predicting the energy efficiency of buildings

For this challenge, you'll work with the renowned UCI Energy Efficiency dataset - real data from 768 building simulations that explores how design choices impact heating and cooling needs. Created by researchers at the University of California, Irvine, this dataset is a machine learning classic that perfectly bridges academic learning with real-world impact.

Goal:

Discover which building features drive energy consumption, compare different prediction models, and translate your findings into concrete business recommendations.

Data Source:

- **Kaggle Dataset:** <https://www.kaggle.com/datasets/elikplim/energy-efficiency-dataset>
 - **UCI Repository:** <https://archive.ics.uci.edu/dataset/242/energy+efficiency>
-

Part 1: Data Exploration & Preparation

Tasks:

1. **Load and explore** the UCI Energy Efficiency dataset
2. **Analyze relationships** between building features (surface area, glazing, orientation) and energy loads
3. **Prepare data** for modeling (train/test split, feature selection)

Ask yourself the following questions:

- Which building characteristics seem to most strongly affect heating/cooling energy consumption?
- Which features could be used for prediction?

Part 2: Model Training & Evaluation

Tasks:

1. **Train at least 2 models** e.g. Linear Regression and Random Forest
2. **Compare performance** using R^2 and RMSE metrics
3. **Analyze feature importance** from both models

Key Questions:

- Which model performs better and why?
- What are the 3 most important features for energy prediction?
- How accurate are your predictions in practical terms?

Part 3: Business Case Documentation

Tasks:

1. **Explore the [appliedAI Institute's Use Case Platform](#)** to get inspired by best practice - this platform represents Europe's largest openly accessible source of curated high-quality AI use cases
2. **Document your own findings** in this project using the [appliedAI Use Case Platform template](#).
3. **Take screenshots** of your completed Use Case template on the platform and add these to a dedicated folder `"/use_case_documentation/"` in your GitHub

Minimal Required Sections for the appliedAI Institute Use Case:

- **Brief description:** What business problem does this solve?
- **Industry:**
- **Value gain:**
- **AI capabilities:**
- **Data sources:**
- **Expected business impact:**

Part 4: Check your work and submit

1. Check your work against the success criteria:

- ✓ R² score > 0.8 on test data
- ✓ Clear business recommendations with quantified impact
- ✓ Complete use case documentation following appliedAI template
- ✓ Professional GitHub repository with organized folder structure

2. Prepare your submission in a .doc or .pdf document which includes:

- Your full name
- The email you use for Kiron's Thrive program
- A bulleted list of specific business recommendations e.g. *"X-facing glazing should be minimized (save €X/year per 100m²)"*
- A working link to your Jupyter notebook for this project via your GitHub repository
 - Please note that the screenshots of your use case template should be included via a dedicated GitHub folder called `"/use_case_documentation/"`

3. Upload your .doc or .pdf document to the 360L platform (final step in AAI-9 course) **by midnight on 21st September**. (Please contact us at thrive@kiron.ngo if you require a one- or two-week extension).