Semantic Mapping of FWH Building Data

P23: Yujin Kim, ykim68

1 Data Set

Our project will use HVAC operation data from the Fitts-Woolard Hall building on Centennial Campus. It contains one year of logsfrom more than 15sensors and equipment points, collected at 15-minute intervals.

2 Project Idea

Our Project goal is to build an auto-labeling model that maps building operation data into a standard format. In Fitts-Woolard Hall, sensor names are inconsistent, so machines cannot understand their meaning (e.g., "AHU1_Temp_S1" is an air handling unit sensor). This inconsistency makes it hard to use building data with other tecnologies or platforms.

We will implement and compare the performance of machine learning models for mapping building points to the Brick Ontology. The comparison will focus on how accurately and consistently each approach can classify diverse building data points.

3 Software to Write

We will write a Python tool to preprocess building data, extract features, and implement an XGBoost-based auto-labeling model that outputs Brick-compatible formats. We will evaluate the model with cross-validation using metrics such as macro F1-socre, precision, and recall.

4 Relevant Papers

[1] Xu, M., Wen, Z., Xiang, X., Fu, W., & Whang, B. (2025) Winning Brick by Brick with Daily Slices: A 94-Task Unified XGBoost Solution for Brick Schema Classification. In *WWW Companion '25*, April 28–May 2, 2025, Sydney, NSW, Australia.

[2] Zhao, H., Macken, J., Dinendra, L., He, Y., & Yang, R. (2025) Hierarchical Multi-Label Classification of Building Management System Time-Series Data. In *WWW Companion '25*, April 28–May 2, 2025, Sydney, NSW, Australia.

[3] Chan, J.H. (2025) BrickMIR: A Minimal, Imbalance-tuned, and Ratio-based Framework for Brick Metadata Classification. In *The WebConf '25*, April 28–May 2, 2025, ICC Sydney, Sydney, Australia.

5 Division of Work

Yujin Kim. Research relevant papers and design the XGBoost-based auto-labeling model. Implement data prepocessing, training, evaluation, and final report preparation.

6 Midterm Milestone

By the Midterm checkpoint, we will implement data preprocessing and feature extraction for the FWH dataset, and ddevelop a prototype XGBoost model for auto-labeling building points.