Capstone III Problem Statement

Background

Energy consumption is a critical issue today. Inside homes, heating, ventilation, and air conditioning (HVAC) systems are responsible for a high amount of the home's total energy consumption, about 53.9%. However, the energy needed to raise or lower temperature in a home is more than the energy required to maintain it. We'd like to develop a system to generate a short term forecast of indoor temperature. The system aims to reduce the power consumption of any home's HVAC. If we can predict the temperature in the house, we can better control when to use HVAC to raise, lower, or simply maintain temperature.

Problem Statement

The goal is, therefore, to reduce energy consumption by predicting the indoor temperature of a room, in order to choose whether or not to activate the HVAC system.

Criteria For Success

We will consider the project successful if we can achieve a root mean squared error of 0.05 or less. In other words, the square root of the average squared distance of our predicted temperatures and the holdout test set's temperatures should be 0.05 C or less. There are no rules for what constitutes an acceptable rmse, but 0.05 is an (arbitrarily) low number.

Scope of Solution Space

Our dataset has 2764 sets of measurements, and sixteen kinds of measurements, not including the target variable temperature. The house the measurements took place in was in the University CEU Cardenal Herrera, just north of Seville, Spain. The predictive model should be appropriate for any house anywhere, as the kind of measurements taken, such as carbon dioxide, humidity, and luminosity, apply to any house.

Constraints

The dataset deliberately has omitted useful features like temperature readings from outside the home. The readings only occurred in the month of March and beginning of April from a location in the northern hemisphere, so the model may predict other seasonal temperatures with less accuracy.

Data Sources

The data comes from a study conducted at the University of CEU Cardenal Herrera (CEU-UCH), Spain, just north of Seville. The dataset can be found at kaggle, and the published

research paper the data was originally gathered for can be found here. Permissions are required to view the paper.