Review of current place machine learning occupies within the field of architecture

* + Define machine learning:
    - Supervised vs unsupervised
  + Tools available: Tensorflow, SciKitLearn, Lunchbox ML
    - Used via python libraries, grasshopper within rhinoceros
* Difference between lunchbox ML grasshopper components, potential uses and examples
* Pros/cons of pure code vs grasshopper for architects/firms:
  + Code allows flexibility, OS agnostic
  + Grasshopper implementation easier to reproduce, teach to coworkers
* Explanation of thesis direction
  + Shift from geometry to energy performance
  + How machine learning can replace traditional building simulation tools with large enough data sets for the algorithm to train on
    - Surrogate modeling
* Possibility of using crowdfunding/crowdsourcing of training data?
  + Even viable to expect designers/firms to contribute their data to an open-source project such as this?

A picture containing diagram

Description automatically generatedFigure 1: Non-linear regression using Lunchbox ML

Artificial intelligence, or AI, has become a standard science fiction movie trope at this point, leading to many misconceptions and in some cases, even fear. Put simply, artificial intelligence is development of computer programs to complete tasks normally requiring human input. One leading field of research within AI is machine learning (ML). Machine learning is a subset of AI in the same way that a square is a quadrilateral but not necessarily vice versa. Machine learning and artificial intelligence have found vastly increasing uses throughout every facet of human life, regardless of whether we, as individuals, notice its impact or not. However, this ubiquity has not lended itself to the use of learning algorithms within the field of architecture. Design tends to be a heavily conservative profession when it comes to the adoption of emerging methods and technologies, but architecture has consistently had an avant-garde wing of researchers and practitioners ready to take the leap.

The vast majority of existing ML research and development has been in two contrasting methods: supervised and unsupervised learning. Both techniques involve feeding vast datasets into a machine learning model in order to ‘train’ the algorithm to accomplish a relatively simple (for a human) task. Supervised learning is built upon training the model using explicitly labeled data in order to predict the *classification* of inputs. On the other hand, unsupervised learning is a means of training a machine learning model using entirely unlabeled data in order to discover insights and/or hidden patterns in the data.

The “hello world!” equivalent, or initial learning exercise for many data science students when first introduced to machine learning consists of training an algorithm to identify hand written digits (0-9). This is the classic example of a supervised learning training method; however, beyond classification lies methods of regression as well. Regression is a method to define a mathematical equation that defines the general alignment of data. These methods range from as simple as linear regression, a slope determining function taught in most high school mathematics curriculums, all the way up to the much more rigorous K-nearest neighbor- a means of organizing data spatially by association.