**Machine Learning Definition**

**Supervised Learning** – Uses historical and labeled data, the ML model predicts a value

* Required historical data
  + Known results and data from the past
* Labeled
  + The desired output is known
* Two main labels typed
  + Categorical Value to Predict
    - Classification Task
      * Predict an assigned category
        + Cancerous vs benign tumor
        + Fulfillment vs credit default
        + Handwriting Recognition
  + Continuous Value to Predict
    - Regression Task
      * Future Prices
      * Electricity loads
      * Test Scores

**Unsupervised Learning** – Uses unlabeled data, the ML model discovers possible patterns in the data.

* Group and interpret data without a label
  + Clustering customers into separate groups based off their behavior
* Major downside is no historical data

**Supervised Learning**

Predicting an Outcome

**Process**

**Selling a home data chart has Area Bedrooms Bathrooms and Price**

**Features-Area, Bedrooms, Bathrooms**

**Label- Price**

Data-> X: Features Y: Label-> Training Data Set (Split) 70% -> Test Data Set (Split) 30% -> Fit/Train Model -> Evaluate Performance -> Deploy Model

**Linear Regression 3.1**

Y = mx + b – slope of a line

**Ordinary Least Squares** – works by minimizing the sum of the squares of the differences between the observed dependent variable (values of the variable being observed) in the given dataset and those predicted by the linear function.

**Mean Absolute Error** – Mean (avg) of the absolute value of the errors. This may miss one or two values where there is a big difference while the majority of the values are close. Larger errors are not “punished” as much

**Mean Squared Error** – Larger errors are “punished” more than with MAE making MSE more popular. Squaring it will make the issue larger. Issue is differences are by the units being squared.

**Root Mean Square Error** – This is the most popular and will take the square root of the values after the units are resolved. This has the same units as y.

**Polynomial Regression**

* Creates the bias (value of 1.0)
* Values raised to a power of degree ( x^1 x^2 x^3)
* Interactions between all pairs of features ( x1\*x2, x1\*x3)

Ex.

* A&B: 1, A, B, A^2, AB, B^2
* X=2, y=3: 1, 2, 3, 4, 6, 9

**Non-linear Relationships** – ex is y=log(x) <- not linear.

**Sync with multiple features** – creating a new feature by multiplying two together.