**Machine Learning Algorithm Notes**

**Linear Regression:**

* One of the most popular algorithms
* Can be used for *classification* and *regression*
* Core algorithms for many other algorithms
  + Logistic, SVM, and modeling decision tree
* Composed of x (IV) and y (DV)
* ***Models*** are lines/(hyper)surfaces in ***feature space*** (IVs and DVs)
  + y = mx + b
  + y = mx + na + pc + d
* R^2 = Mean Squared Error
  + Way of judging the *goodness of fit*
* Examples
  + Temperature vs. Amount of ice cream sold
  + Inches of rain vs. new cars sold
  + Daily snowfalls vs. number of skiers
* If you suspect a relationship/correlation between two variables, a regression can help confirm it
  + R^2 is a numerical way of determining model performance

**Logistic Regression:**

* Useful where there are two distinct y-values

**Linear vs. Logistic Regression**

|  |  |
| --- | --- |
| **Linear** | **Logistic** |
| Continuous Variables | Categorical Variables |
| Regression Solution | Classification Solution |
| Straight line | S-curve |



* When deciding a model/algorithm, look at what output you are looking for

**A**