Understanding Machine Learning

Predictive Analytics

Predictive Analytics

Predictive analytics attempts to determine an unknown variable based on known variables using probabilities.
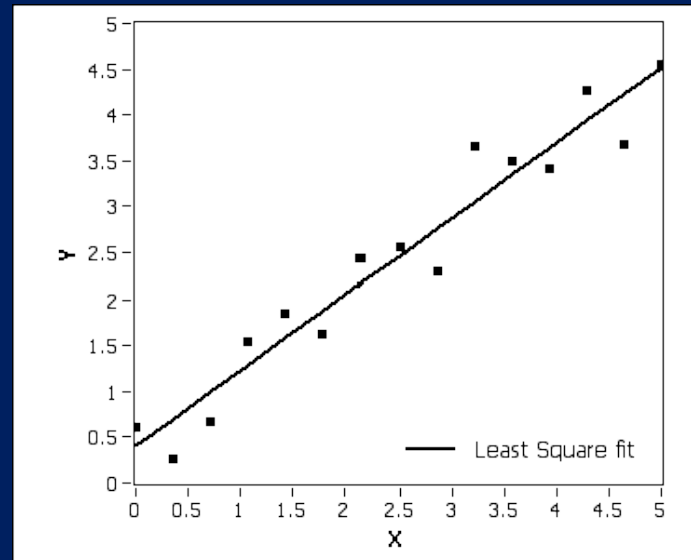
Predictive Analytics

History can be used to predict the future.

- Use historical data to predict an outcome.

- Need to have historical data with both the outcome and other strongly correlated columns called features.

- Model training calculates the statistical probabilities of features value combinations with the outcome.

What is a Predictive Model?

Determines the probable value of some variable based on other variables.

Example: Gross Sales related to Profit, i.e. the more sales revenue, generally the more profit. Assume Y is Sales and X is profit over a series of years.

The points are the actual sales/profit. The line is attempting to get the average. This is an example of a linear relationship.

Other Linear Relationships

- Foot size to height.

- Study time to grades.

- Distance to travel time.

Relationship is calculated as a number between 0 and 1 with 0 being no relationship and 1 being a 1 for 1 relationship.

A negative value indicates an inverse Relationship, i.e. x goes up, y goes down.
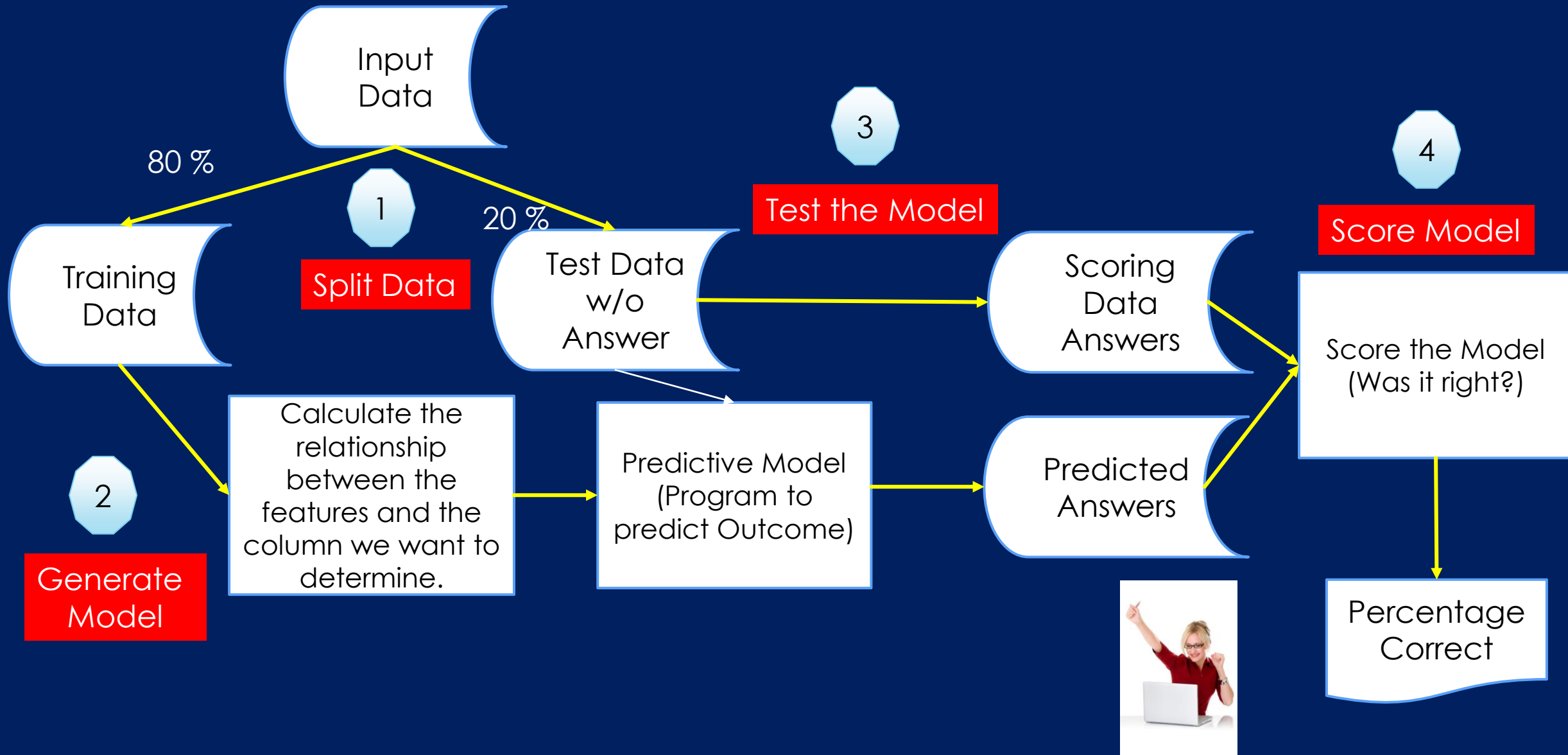
Uses of Predictive Models

- Classify something such as is an is an email spam or ham?

- What group does something belong to such as what species of plant?

- Will a person contract a certain illness?

- Will the patient be readmitted?

- Is the person a high health risk?

Steps to Creating a Predictive Model

- Analyze/Clean the data.

- Determine the columns that pred

- Select a model to use.

- Train the model.

- Evaluate the model's effectiveness.

- Use the model.

Data Preparation is 80% of Data Science Work!

Steps to Training a Predictive Model

Training a Predictive Model

How long will it take to drive into Boston?

Column	Meaning
Starting Distance	How far are you from Boston?
Day of the Week	**Weekdays are busier**
Is it a Holiday?	May be less traffic
Time of Day	Rush hours are busier
Estimated Time	Number of minutes to get to Boston

Predictive Modeling – Wrapping Up

Questions?