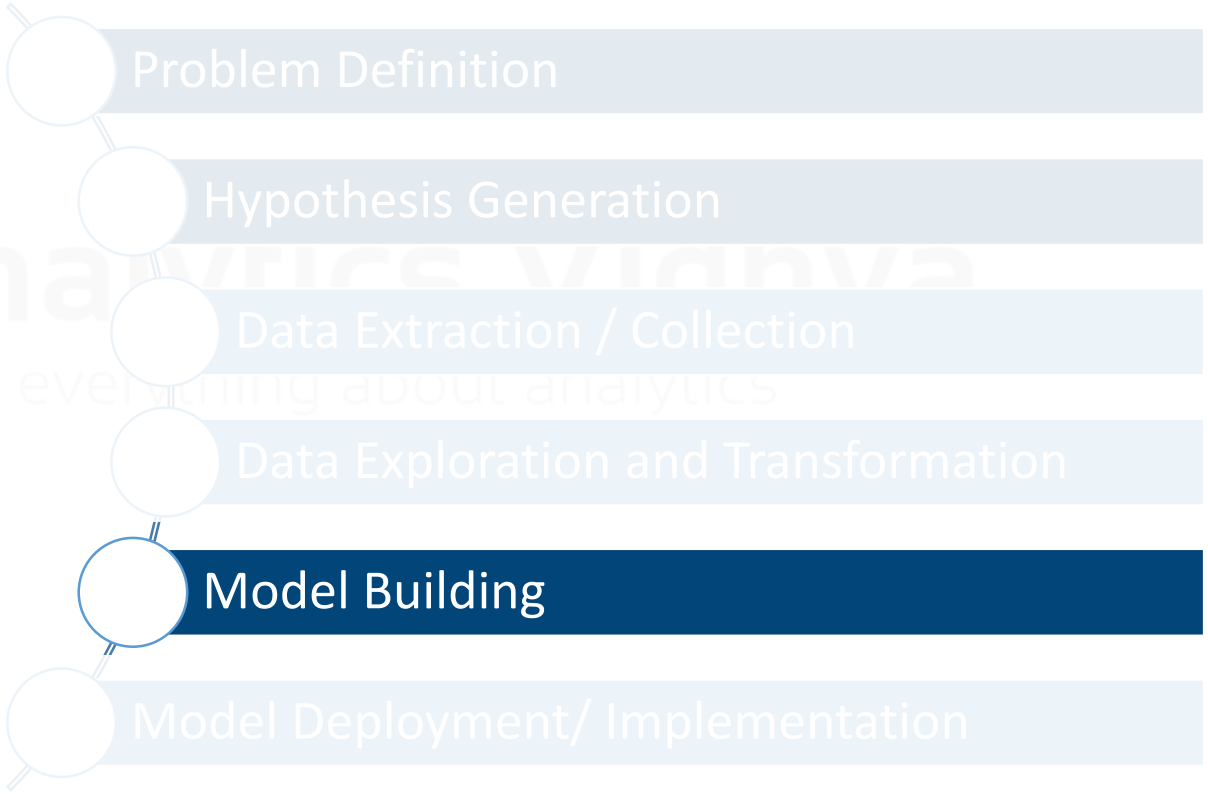
Stages of Predictive Modeling

We can broadly divide the model building life cycle in six stages:



Model Building

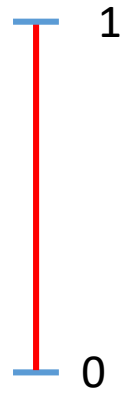
It is a process to create a mathematical model for estimating/ predicting the future behaviour based on past data.

Model Building

Example:

A retail bank wants to know the default behaviour of its credit card customers. They want to predict the probability of default for each customer with in next 3 months.

- Probability of default would lie between 0 and 1.
- Assume every customer has a 10% default rate.



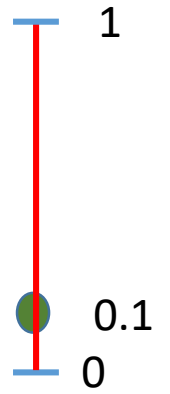
Model Building

Example:

A retail bank wants to know the default behaviour of their credit card customers like “Predict the probability of default for customers?”

Probability of default for each customer in next 3 months = 0.1

Now, what these “Predictive Models” do?

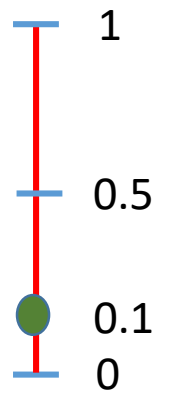


Model Building

Example:

A retail bank wants to know the default behaviour of their credit card customers like “Predict the probability of default for customers?”

Probability of default for each customer in next 3 months = 0.1



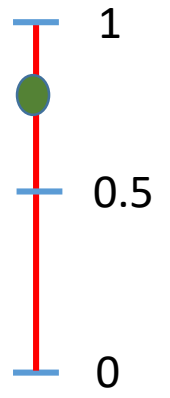
It moves the probability towards one of the two extremes based on attributes from past information

Model Building

Example:

A retail bank wants to know the default behaviour of their credit card customers like “Predict the probability of default for customers?”

Probability of default for each customer in next 3 months = 0.1



It moves the probability towards one of the two extremes based some given attributes

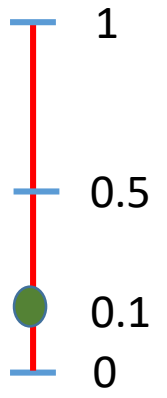
A customer with volatile income is more likely (closer to 1) to default

Model Building

Example:

A retail bank wants to know the default behaviour of their credit card customers like “Predict the probability of default for customers?”

Probability of default for each customer in next 3 months = 0.1



It moves the probability towards one of the two extremes based some given attributes

OR, A customer with healthy credit history for last 10 years has low chances of default (closer to 0)

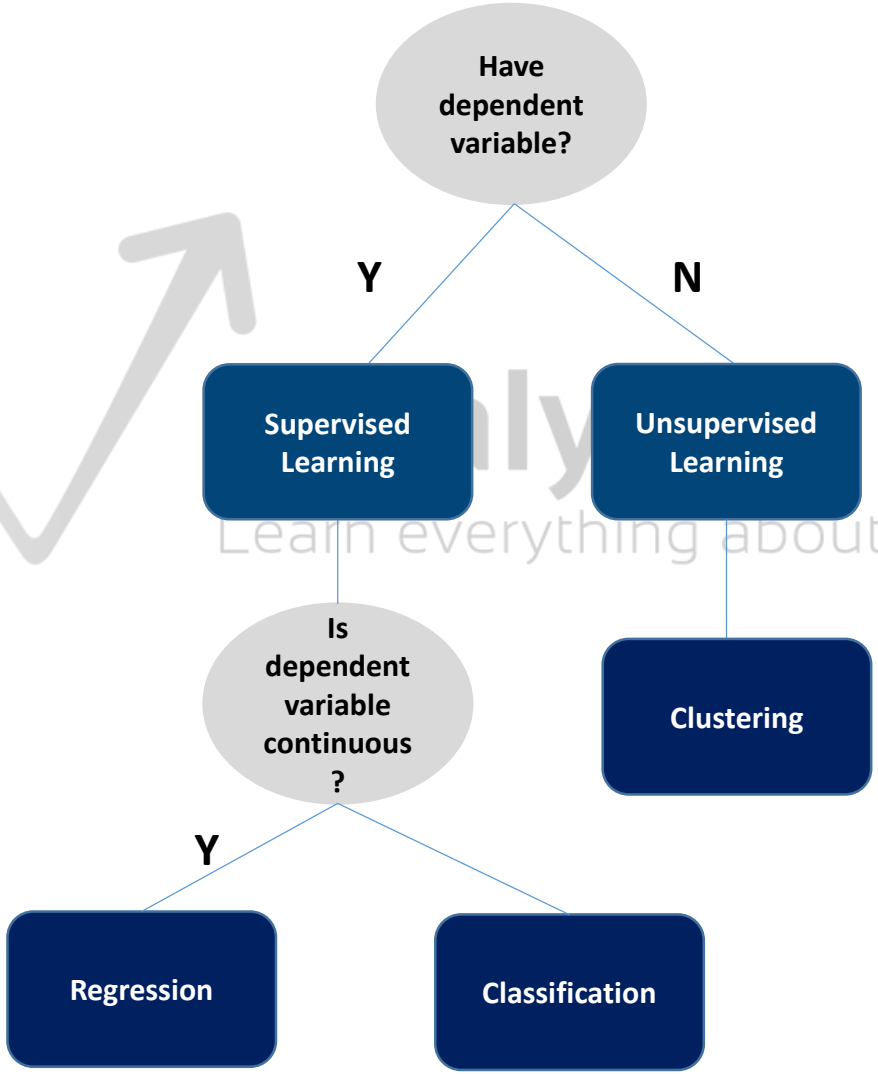
Steps of Model Building

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Example:

Customer segmentation for defining right product

Common Algorithms:

- K-means
- Spectral Clustering

Steps of Model Building

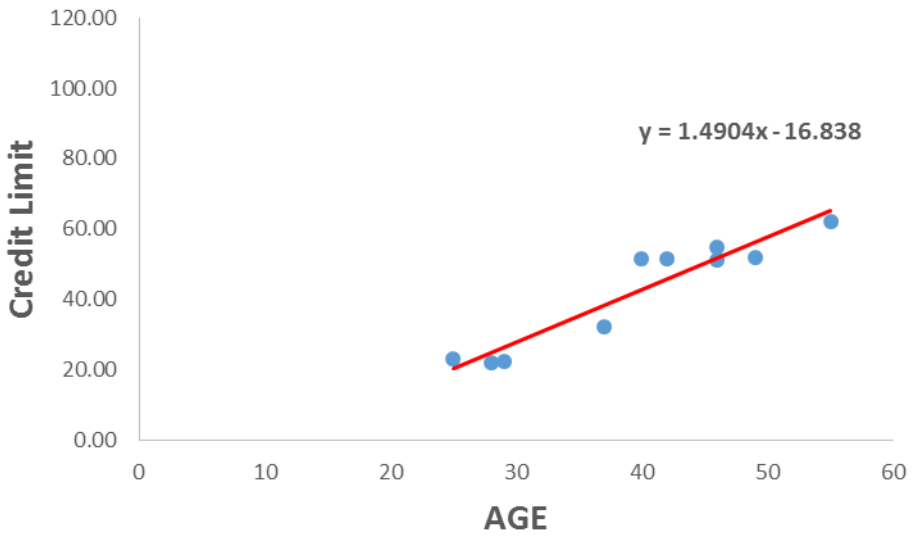
It is a process to learn relationship/ correlation between independent and dependent variables

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

Here, you can see predictive model (linear equation) based on single variable

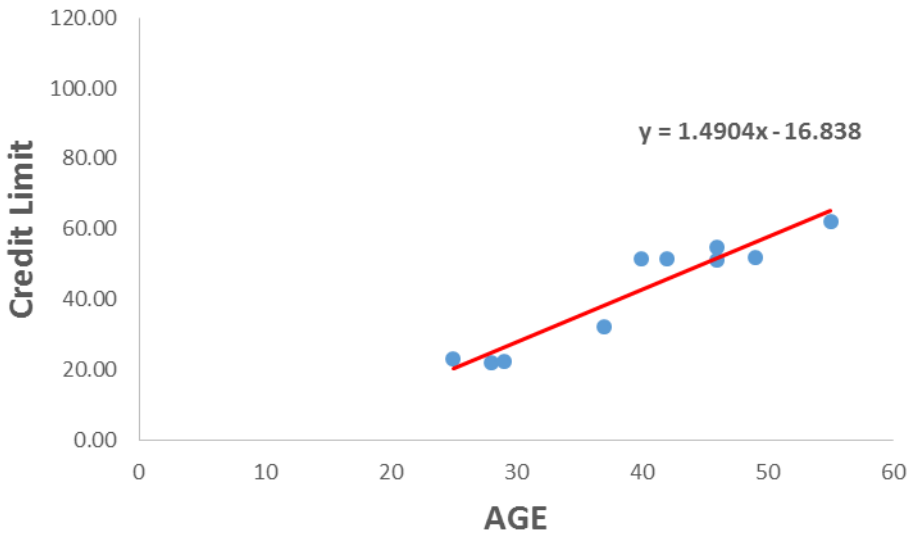
Analytics Vidhya
Learn everything about analytics

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

Where will we use this equation?

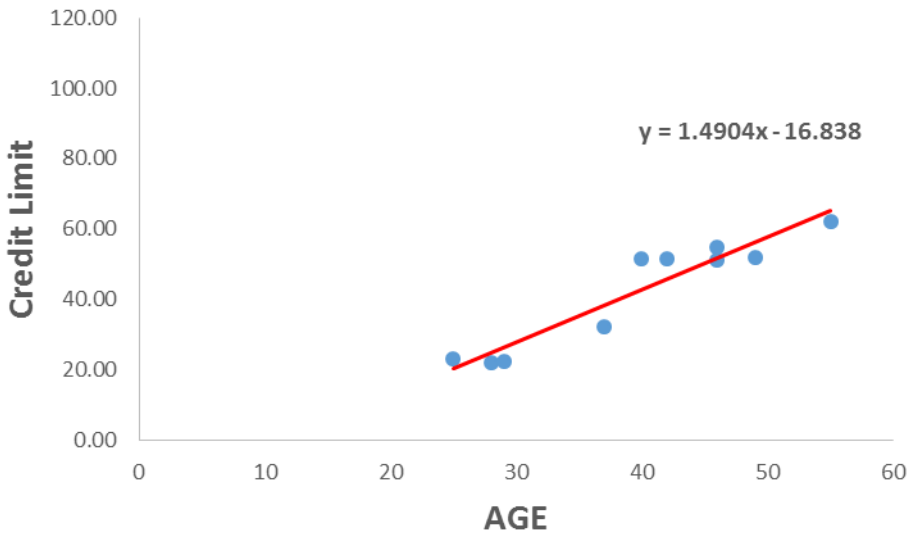
Analytics Vidhya
Learn everything about analytics

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

This equation is used to predict/ estimate dependent variable of test data set

Analytics Vidhya
Learn everything about analytics

Steps of Model Building

Algorithm Selection

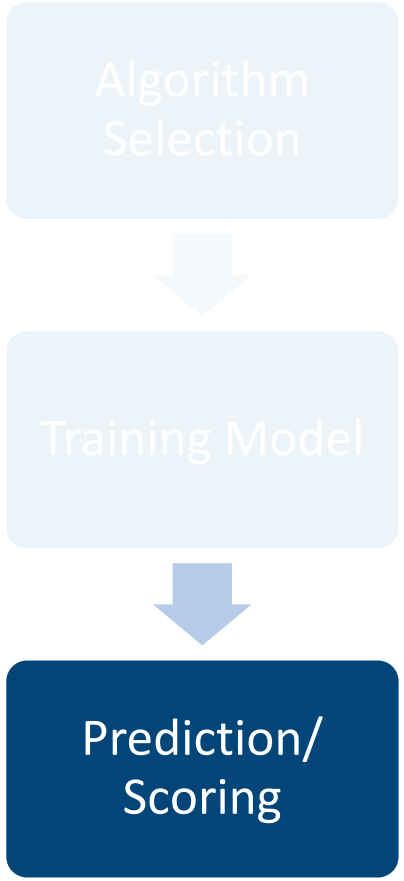
Training Model

Prediction/ Scoring

- Past data (Known dependent variable)
- Used to train model

- Future data (Unknown dependent variable)
- Used to score

Steps of Model Building



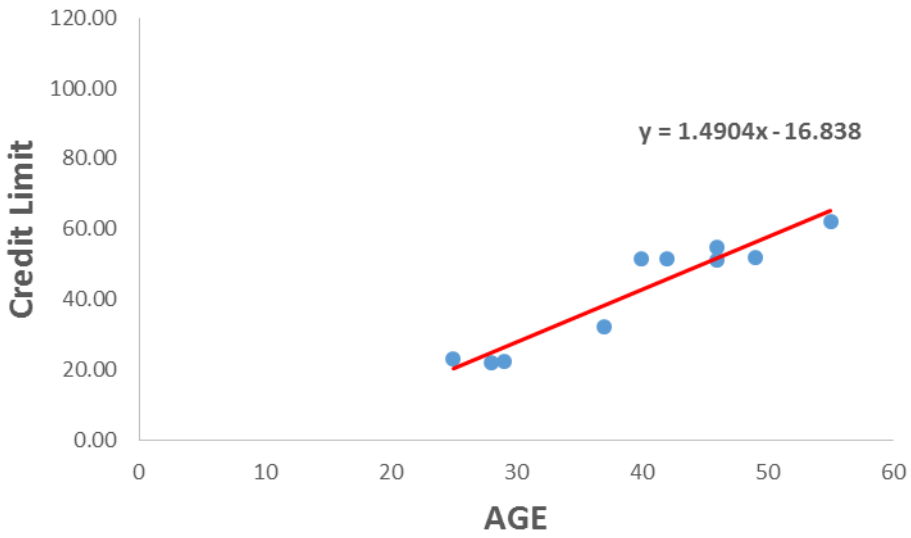
It is a process to estimate/ predict dependent variable of test data set by applying model rules

Steps of Predictive Modeling

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

For “Age” 40, Credit Limit = 1.4904*40 -16.838 = 42.778

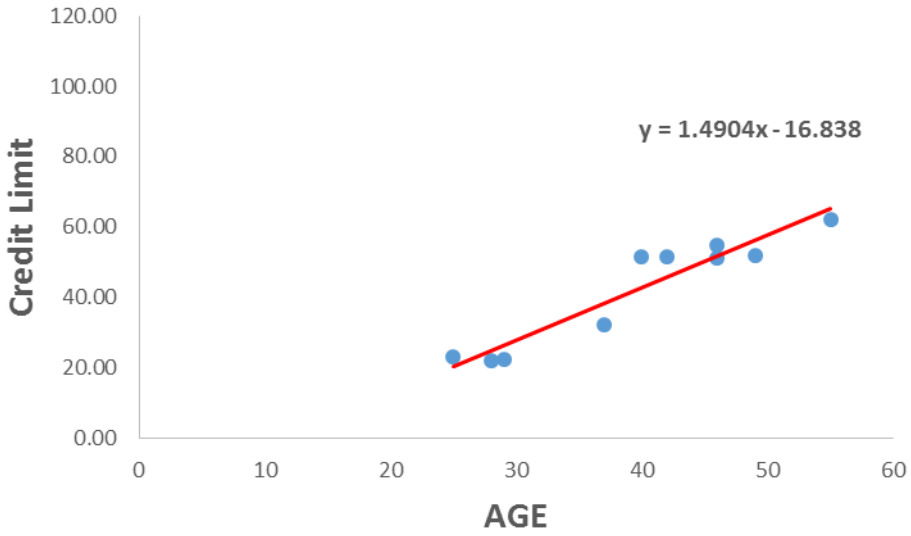
Analytics Vidhya
Learn everything about analytics

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

Similarly, for “Age” 50, Credit Limit = 1.4904*50 -16.838 = 57.682

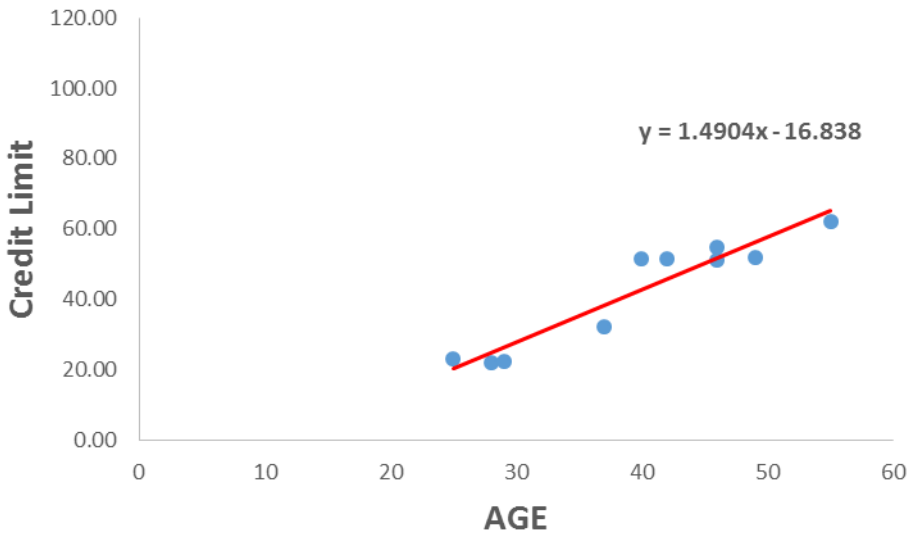
Analytics Vidhya
Learn everything about analytics

Steps of Model Building

Algorithm Selection

Training Model

Prediction/ Scoring



Equation:

Credit Limit = f(Age)

Credit Limit= 1.4904*Age - 16.838

We always apply training learning to test data set for prediction/ estimation

Analytics Vidhya
Learn everything about analytics

Thank You

Analytics Vidhya

Learn everything about analytics

Analytics Vidhya
Learn everything about analytics