



# Machine Learning Introduction

In support of Oracle Machine Learning tutorial

---

**Peter Jeffcock**

Data Science and Big Data Product Marketing

Organization, Division or Business Unit

April 2020

## Program agenda

---

- 1 Machine learning overview
- 2 Classification
- 3 Decision trees

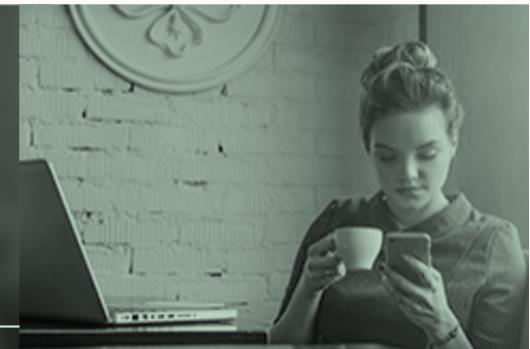


## Have you ever...

Had a credit card transaction unexpectedly (and incorrectly) declined



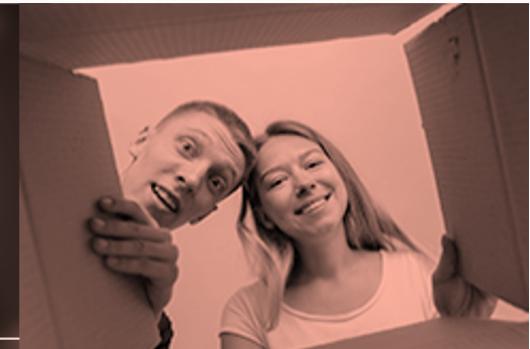
Received a personalized email, direct mail or web ad



Been influenced by a recommendation from your music player or an online shopping site



Had an unexpected bonus / incentive to stay with a company





**Machine Learning** is based on algorithms that can learn from data without relying on rules-based programming.

---

**McKinsey**

# Classifying customers

Basic query	Age/Gender	Known	Known	Known	Unknown	Unknown
	Marketing preferences	Mail, email	Email	Email, Facebook	Email, twitter	Mail, email, Twitter
Basic analytics	RFM: Purchases in last 3/6/12 mo	1 item/\$35 in the last 3 months	2 items/\$150 in the last 6 months	3 items/\$75 in the last 3 months	3 items/\$225 in the last 12 months	9 items/\$350 in the last 6 months
Machine learning	Behavioral segment	Retired cosmopolitan	Affluential executive	New home mom	Successful startup	Executive collector
	Credit	Good	Other	Good	Other	Good



## Classifying fruit

Large basket of fruit

Pass individual fruits around, telling you what they are

Keep 20-30% of the fruit back. Then pass those around to test your understanding. Are you good enough?

Put to work sorting fruit

Bring you back for more training



## Mapping to Machine Learning Terms

Term	We used	Data scientist might use
Technique	Classify (fruit)	Classification, clustering, time series. etc.
Algorithm	“Deep learning neural network” in your brains	Neural network, decision tree, k-means clustering, etc.
Training Data	The initial basket of fruit	Data set supplied by IT, often setup and cleaned up by data scientist
Training the model	You figured it out	Adjust different parameters in response to the data to make it more accurate
Testing the model	The left-over fruit	Always reserve some data that the model hasn't seen to test.
Model deployment	Sent you to fruit packing line	Make model available to app developers, execs, analytics tools etc.
Model update	Brought you back for more training	Build a new model or re-train the old on additional data. Must re-deploy





## Decision Trees

---

## Predicting churn

---

5 attributes

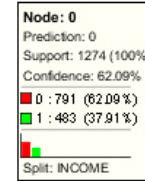
3 predictors

ID	Age	Income	Gender	Churned
1008764	34	47,200	F	Yes

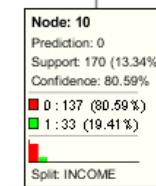
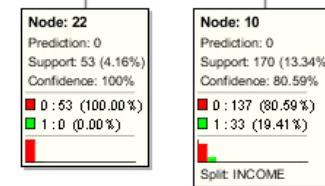
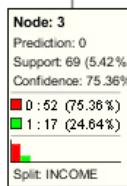
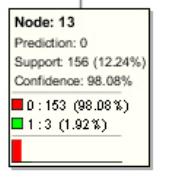
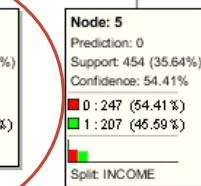
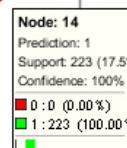
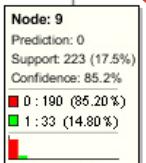
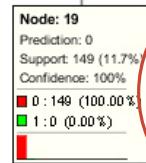
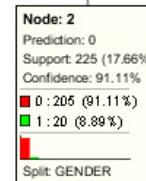
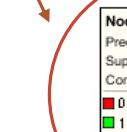


Prediction: Will Churn  
 Confidence: 64%  
 Supported by: 677 (53%)  
 Description:  $41,500 < \text{Income} < 89,500$

Leaf Node



Prediction: Won't Churn  
 Confidence: 85%  
 Supported by: 223 (18%)  
 Description:  $129,500 < \text{Income}$



For more background look at <https://blogs.oracle.com/bigdata/decision-trees-machine-learning>



# Thank you

---





**ORACLE**