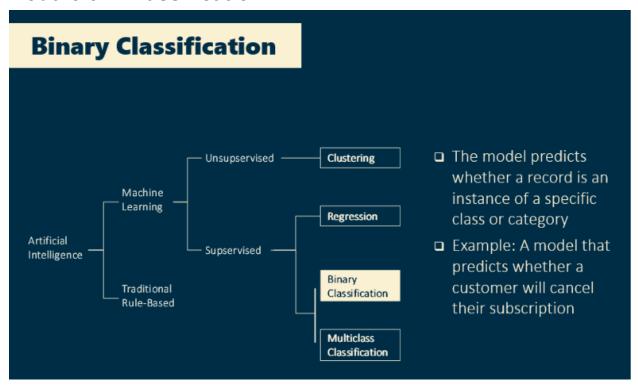
Module 04: Classification



In Supervised Machine Learning, inputs are called **feature** values while outputs are called **label** values

Age	Gender	Booked Hotel	Linked To Social Media	Frequent Flyer Member	Income Class	
24	М	Y	N	Y	Middle Income	
29	М	N	N	N	Middle Income	
34	F	Y	Y	Y	High Income	
43	М	Y	N	N	High Income	
44	F	N	N	N	Middle Income	
27	F	Y	Y	Y	Low Income	
33	М	N	N	N	Middle Income	
31	F	N	Y	N	Low Income	
22	F	Y	N	Y	Low Income	
26	F	N	Y	N	Middle Income	
51	М	N	Y	N	Middle Income	

Feature values

Churn
Y
N
N
N
N
Y
N
Y
Y
N
N



The available data set is split into training data and test data. Typical training data is around 70% to 80%.

Training Data

24	M	Y	N	Y	Middle Income	Y
29	M	N	N	N	Middle Income	N
34	F	Y	Y	Y	High Income	N
43	M	Y	N	N	High Income	N
44	F	N	N	N	Middle Income	N
27	F	Y	Y	Y	Low Income	Y
33	M	N	N	N	Middle Income	N

Test Data

31	F	N	Y	N	Low Income	Y
22	F	Y	N	Y	Low Income	Y
26	F	N	Y	N	Middle Income	N
51	M	N	Y	N	Middle Income	N