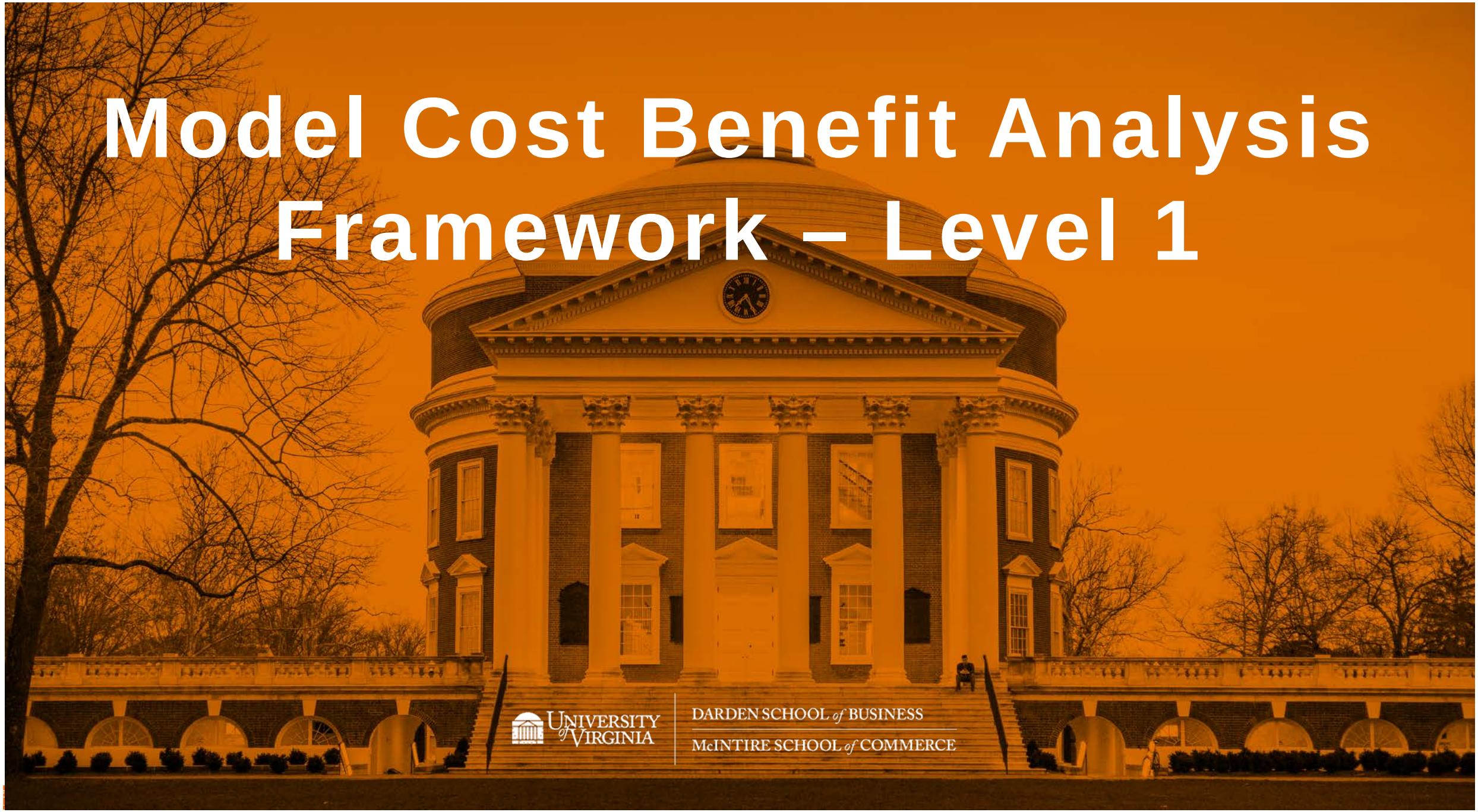
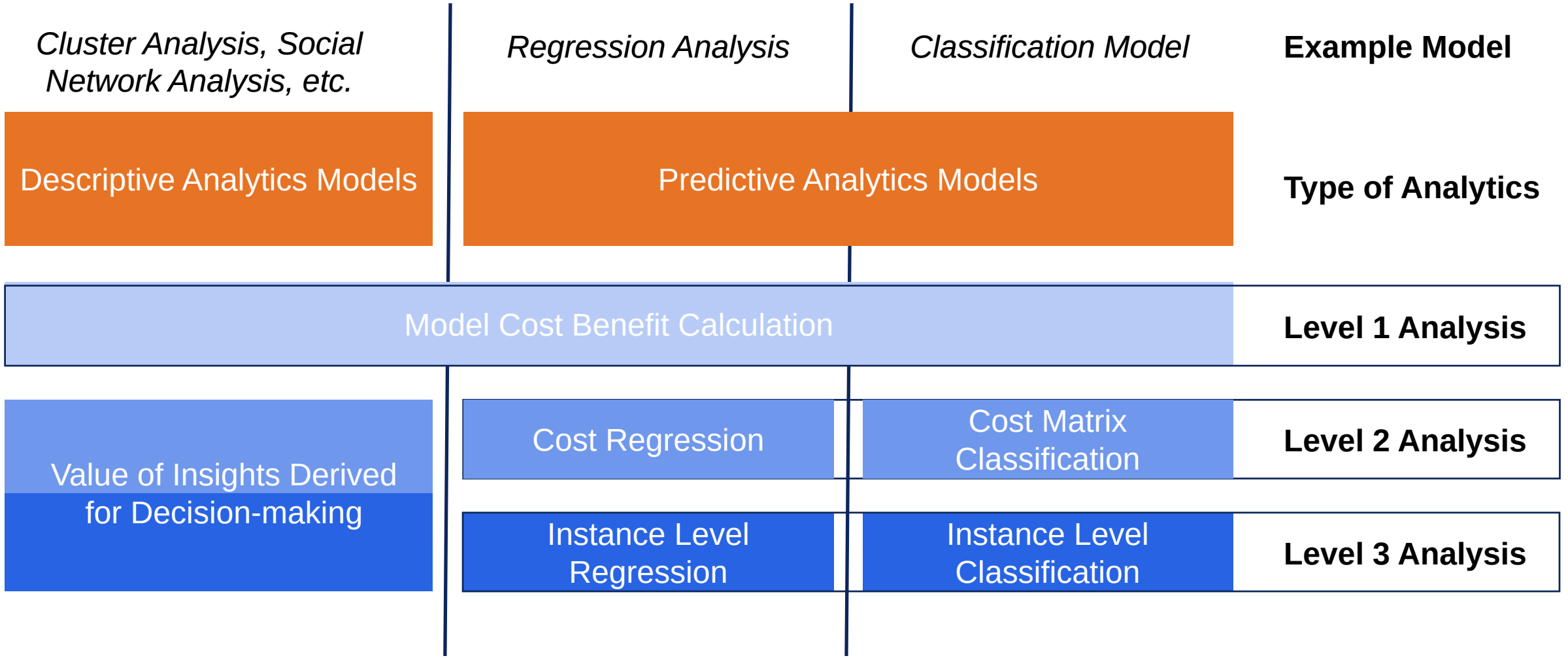
Model Cost Benefit Analysis Framework – Level 1



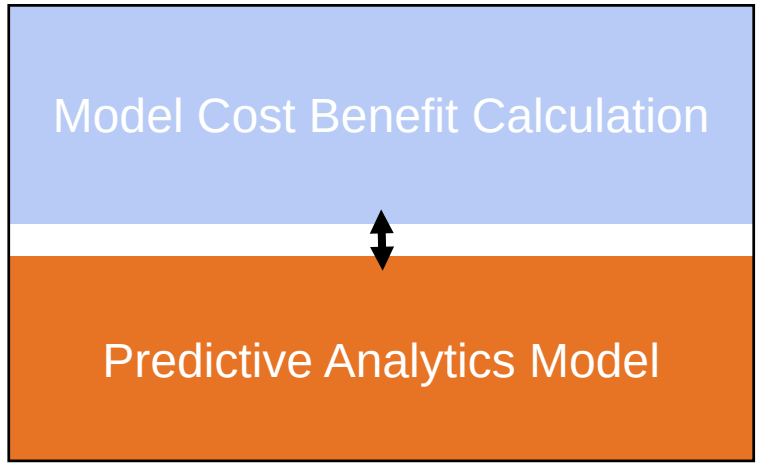
Model Cost-benefit Analysis Framework



Model Cost Benefit Analysis Framework – Predictive Analytics

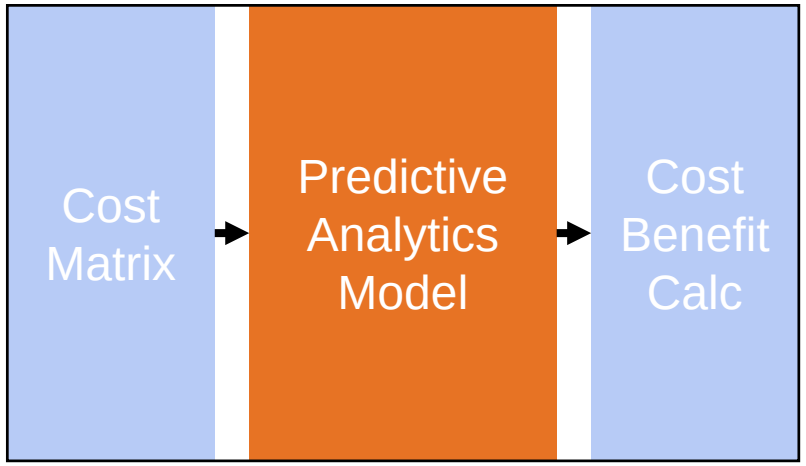
Three potential levels of analysis, depending on the problem context

Lv 1 – Model Level



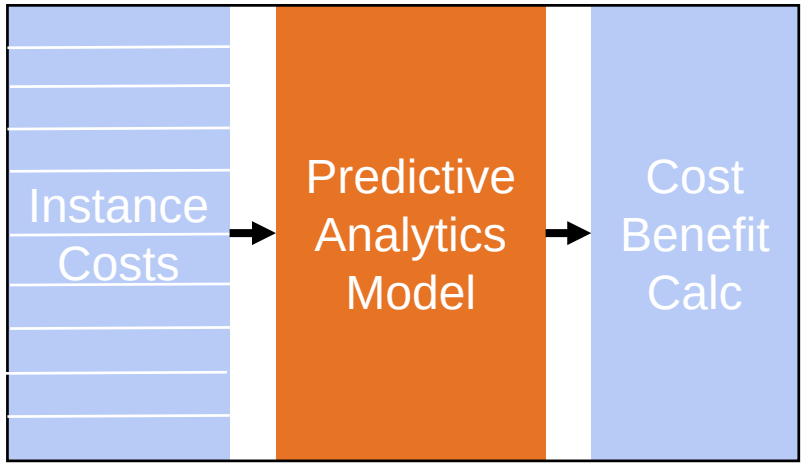
Treat model like a black box
Couple TP/TN/FP/FN with cost-benefit calculations

Lv 2 – Matrix Level



Input cost matrix ratios into model
Couple TP/TN/FP/FN with cost-benefit calculations

Lv 3 – Instance Level

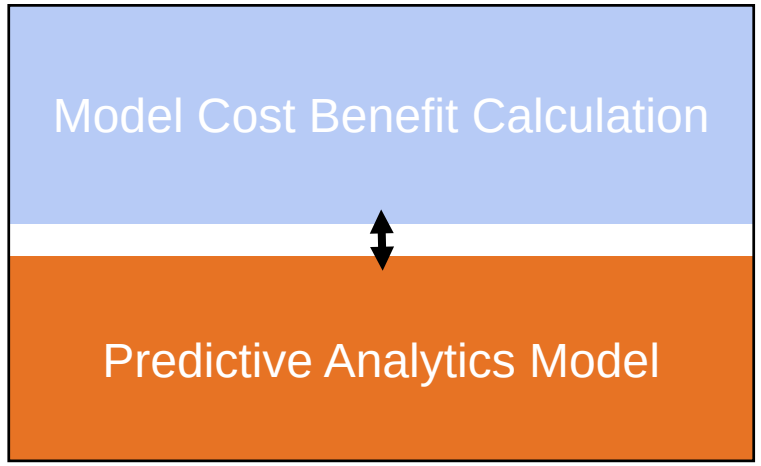


Input instance-level costs into model
Couple TP/TN/FP/FN with cost-benefit calculations

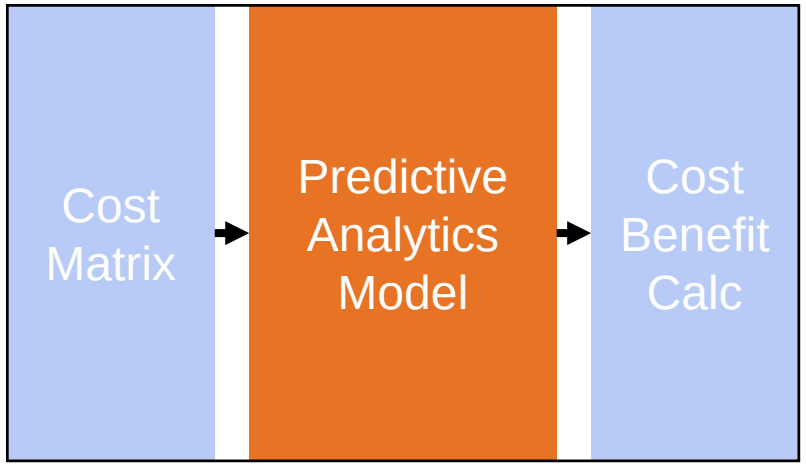
Model Cost Benefit Analysis Framework – Predictive Analytics

Three potential levels of analysis, depending on the problem context

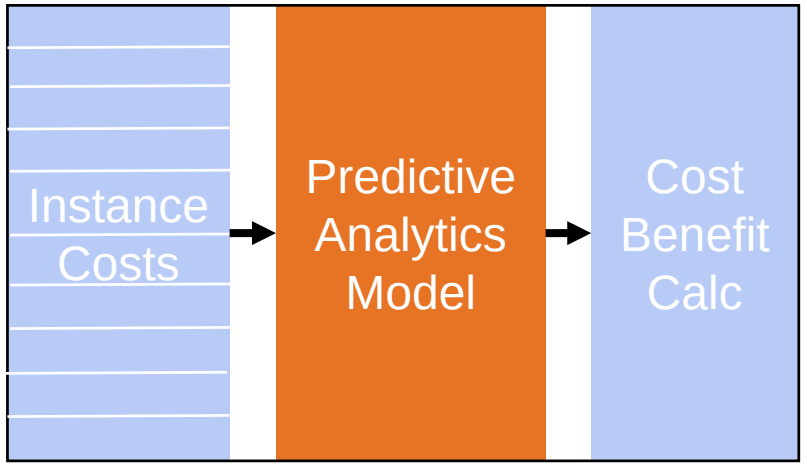
Lv 1 – Model Level

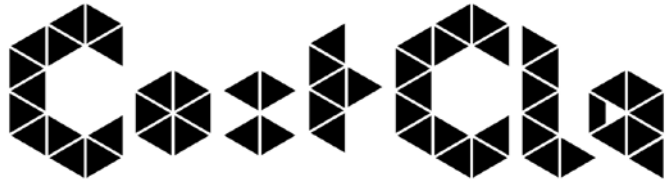


Lv 2 – Matrix Level

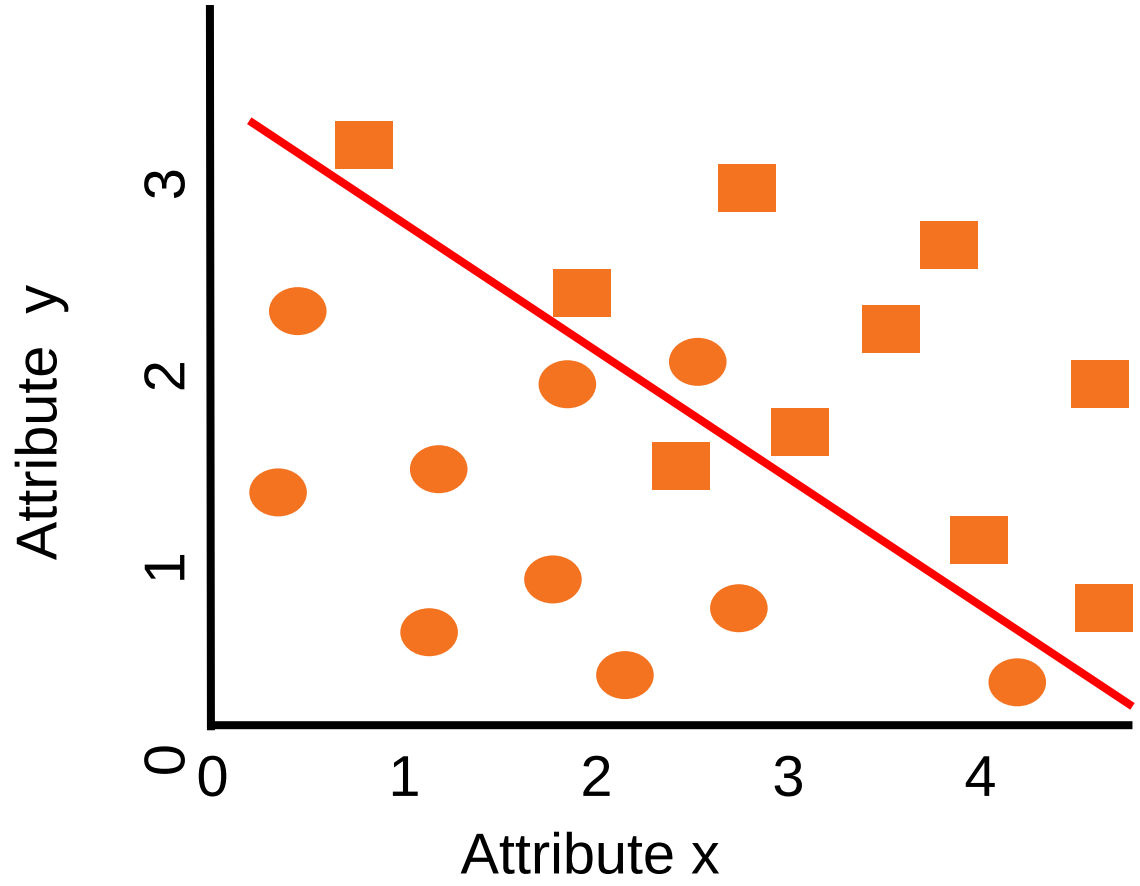


Lv 3 – Instance Level





Model Cost-benefit Analysis Framework – Motivation



Draw a single straight line that can best separate circles from squares (i.e., minimal error rate)

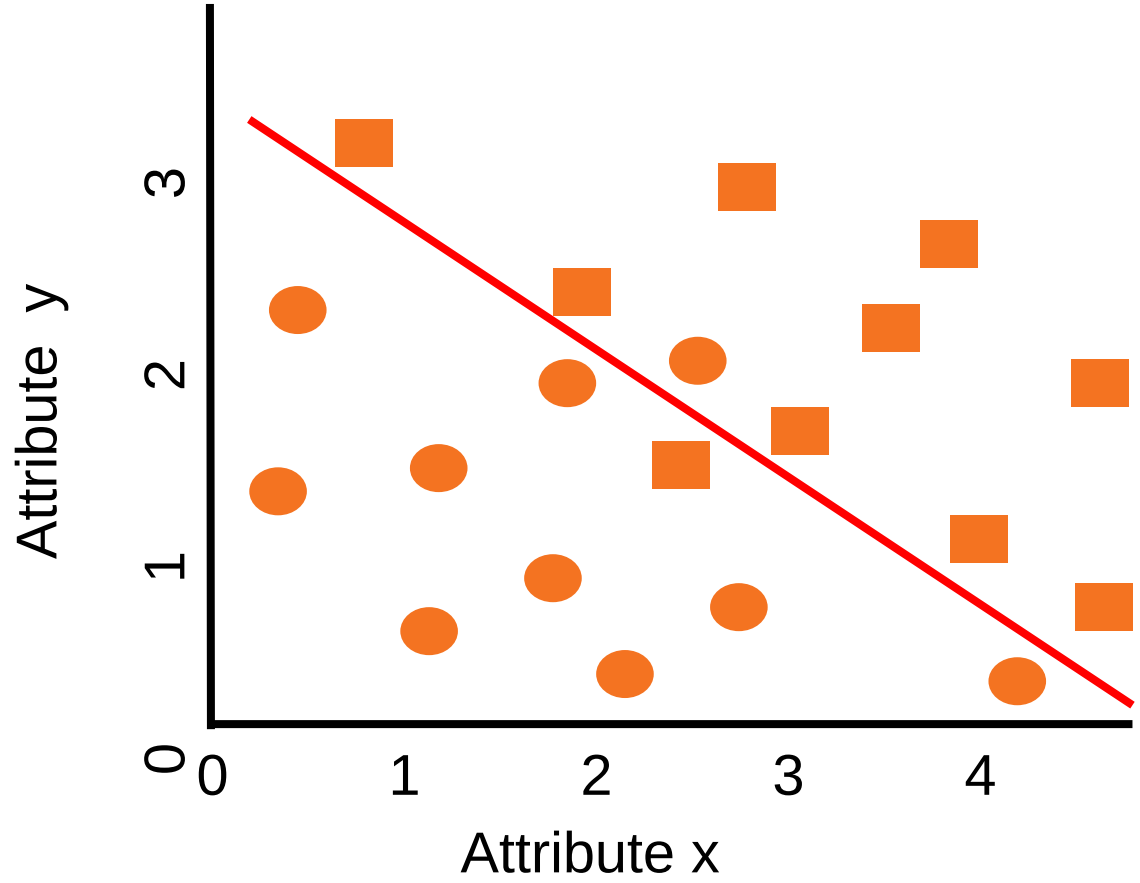
Performance on Training Data:

Accuracy = 18/20 = 0.90

Circle Recall = 9/10 = 0.90

Square Recall = 9/10 = 0.90

Model Cost-benefit Analysis Framework – Equal Costs



Draw a single straight line that can best separate circles from squares (i.e., minimal error rate)

Performance on Training Data:

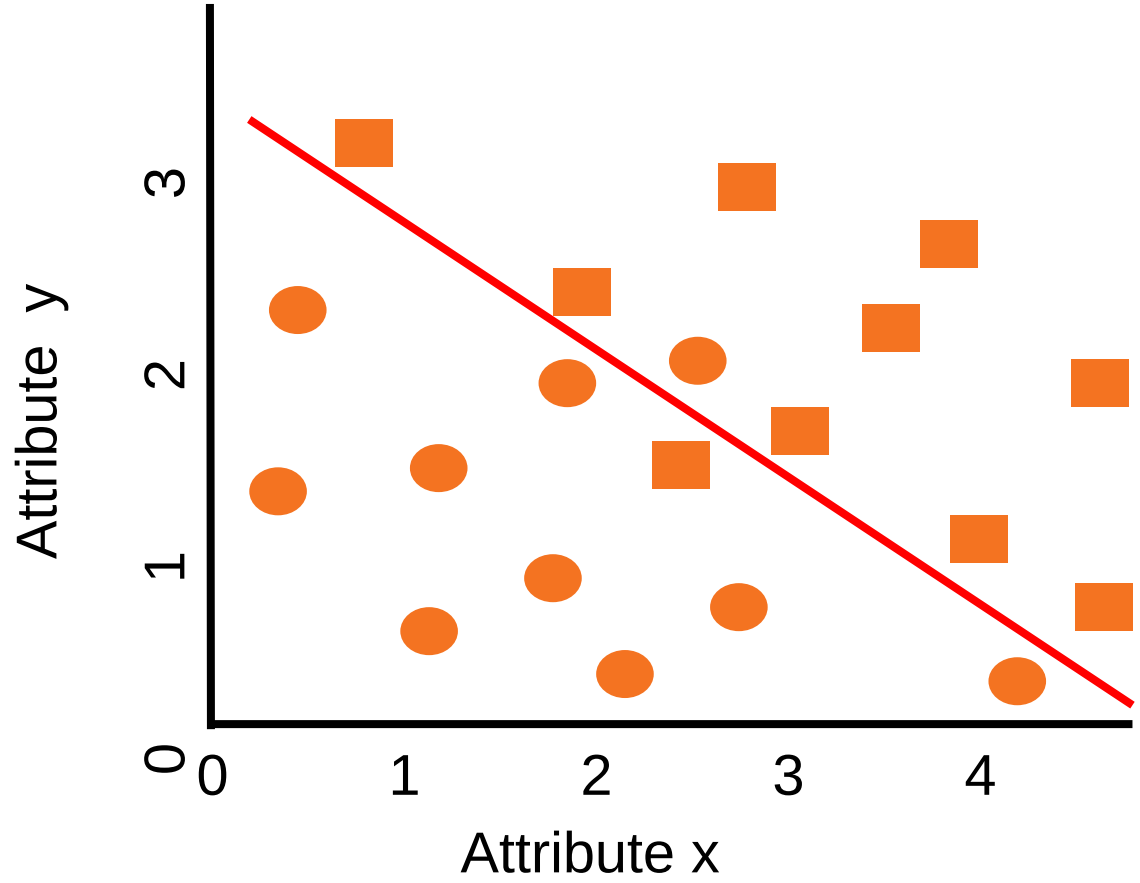
Accuracy = 18/20 = 0.90

Circle Recall = 9/10 = 0.90

Square Recall = 9/10 = 0.90

Cost of False Sq = 1; Cost of False Cr = 1

Model Cost-benefit Analysis Framework – Asymmetric Costs



Draw a single straight line that can best separate circles from squares (i.e., minimal error rate)

Performance on Training Data:

Accuracy = 18/20 = 0.90

Circle Recall = 9/10 = 0.90

Square Recall = 9/10 = 0.90

Cost of False Sq = 3; Cost of False Cr = 5

Model Cost-benefit Analysis Framework – Predictive Analytics Lv 1

Model level analysis example – what we discussed in confusion matrix video

Confusion Matrix

Predicted Class	Actual Class: Class = Yes	Actual Class: Class = No
Class = Yes	TP	FP
Class = No	FN	TN

Cost Matrix

Predicted Class	Actual Class: Class = Yes	Actual Class: Class = No
Class = Yes	TP cost	FP cost
Class = No	FN cost	TN cost

Model Cost-benefit Analysis Framework – Predictive Analytics Lv 1

Model level analysis example

Confusion Matrix

Predicted Class	Actual Class: Class = Yes	Actual Class: Class = No
Class = Yes	TP	FP
Class = No	FN	TN

Cost Matrix

Predicted Class	Actual Class: Class = Yes	Actual Class: Class = No
Class = Yes	TP cost	**FP cost**
Class = No	**FN cost**	TN cost

Step 1: Figure out cost matrix values

Step 2: Typically focus on FP/FN costs. *Model value* is cost reduction over some status quo

Step 3: Total cost = (FP cost x FP rate x Total annual instances)

+

(FN cost x FN rate x Total annual instances)

Model Cost-benefit Analysis Framework – Predictive Analytics Lv 1

Model level analysis example – fraud detection illustration assuming 1000 total cases per year

A

Predicted Class	Actual Class		
	Class = Yes	Class = No	**Precision**
Class = Yes	3	1	**75%**
Class = No	2	4	**66.7%**
	Recall	**60%**	**80%**

Step 3 for Model A: Total cost = ($500 x 10% x 1000) = **$50,000**

+

($1000 x 20% x 1000) = **$200,000**

B

Predicted Class	Actual Class		
	Class = Yes	Class = No	**Precision**
Class = Yes	2	0	**100%**
Class = No	3	5	**37.5%**
	Recall	**40%**	**100%**

Step 3 for Model B: Total cost = ($500 x 0% x 1000) = **$0**

+

($1000 x 30% x 1000) = **$300,000**

Cost of audit is $500;
Cost of fraud is $1000

?


VIRGINIA