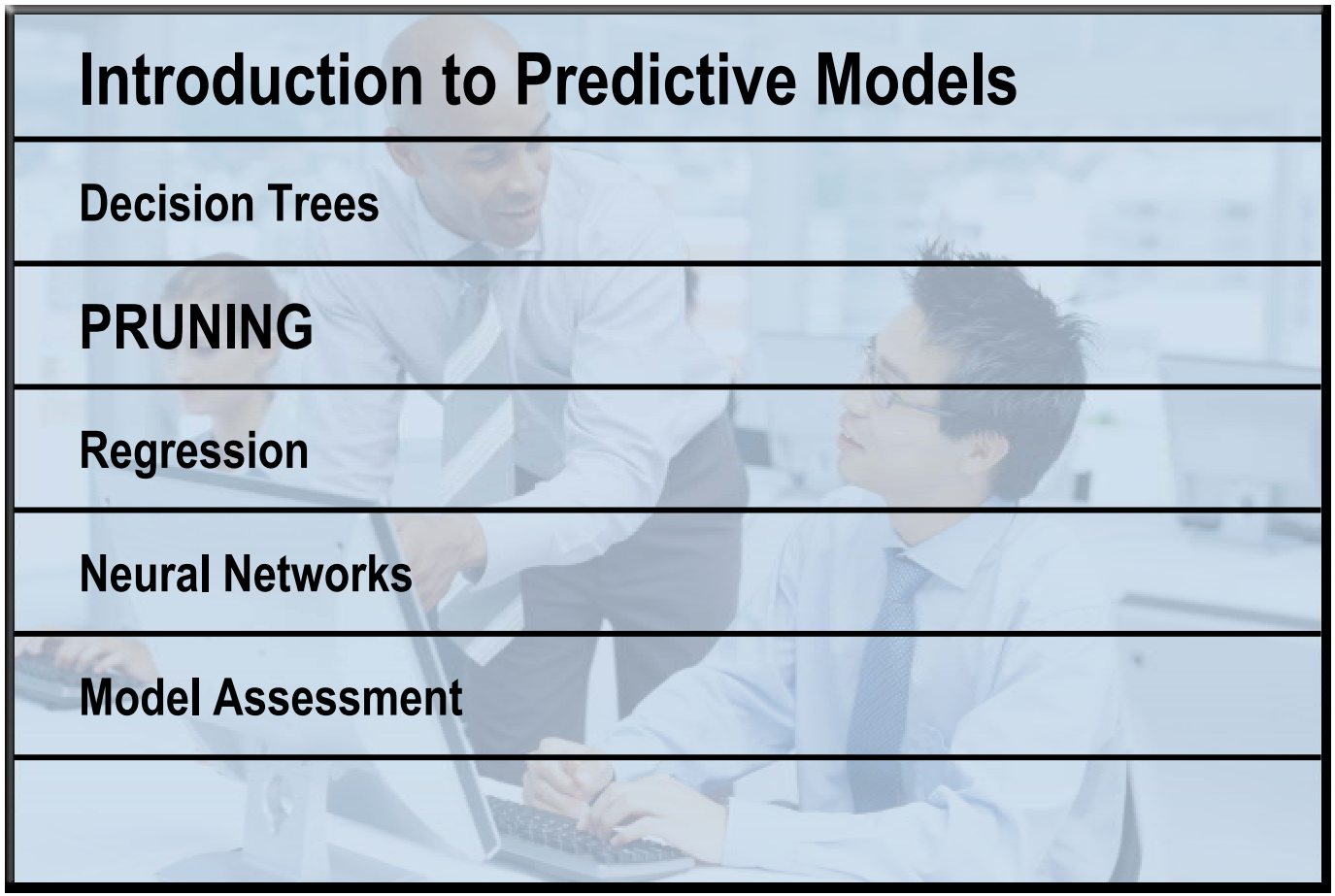
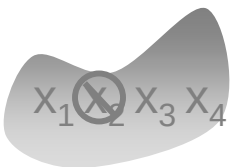
Agenda


Introduction to Predictive Models
Decision Trees
PRUNING
Regression
Neural Networks
Model Assessment

Model Essentials – Decision Trees

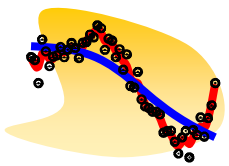
Predict new cases

Prediction rules



Select useful inputs

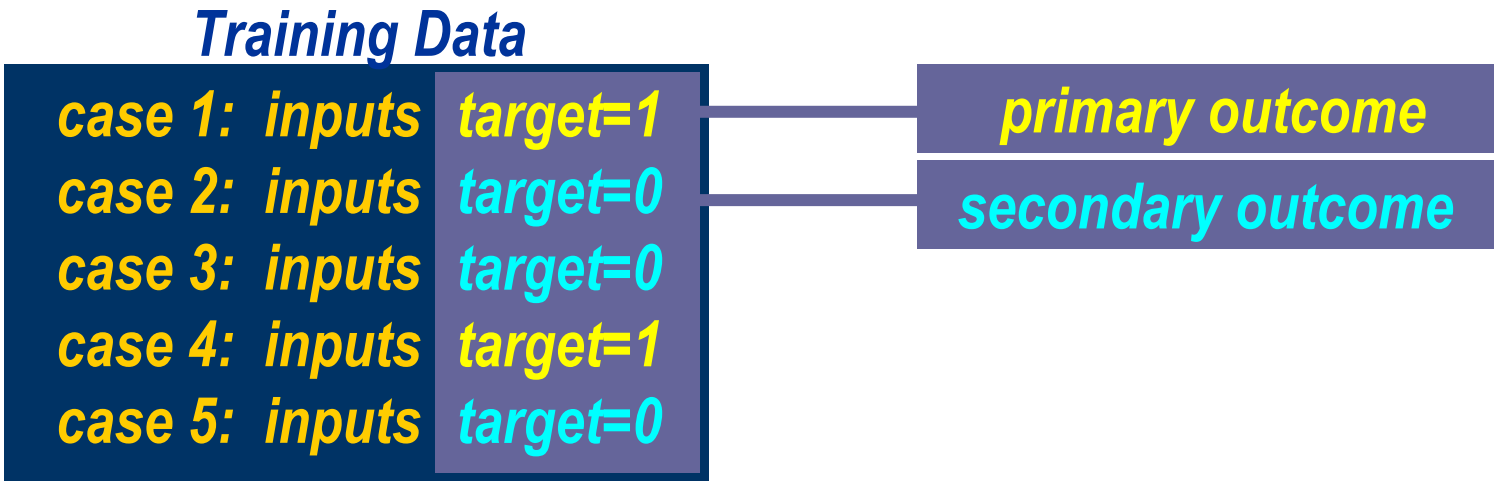
Split search



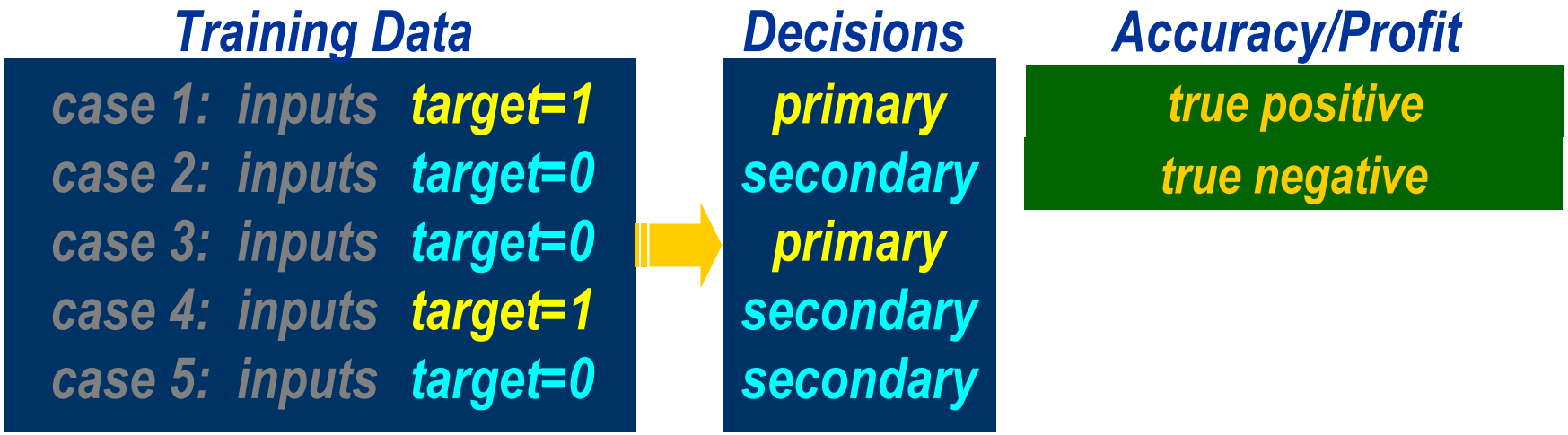
Optimize complexity

Pruning

Binary Targets

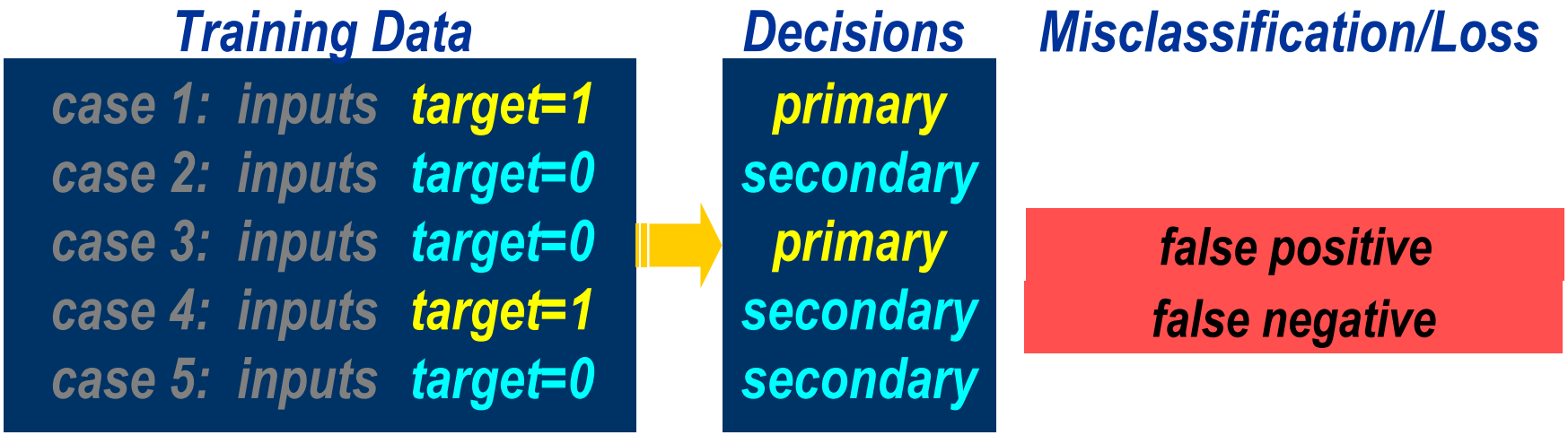


Decision Assessment



Focus on correct decisions

Decision Assessment (for Pessimists)



Focus on incorrect decisions

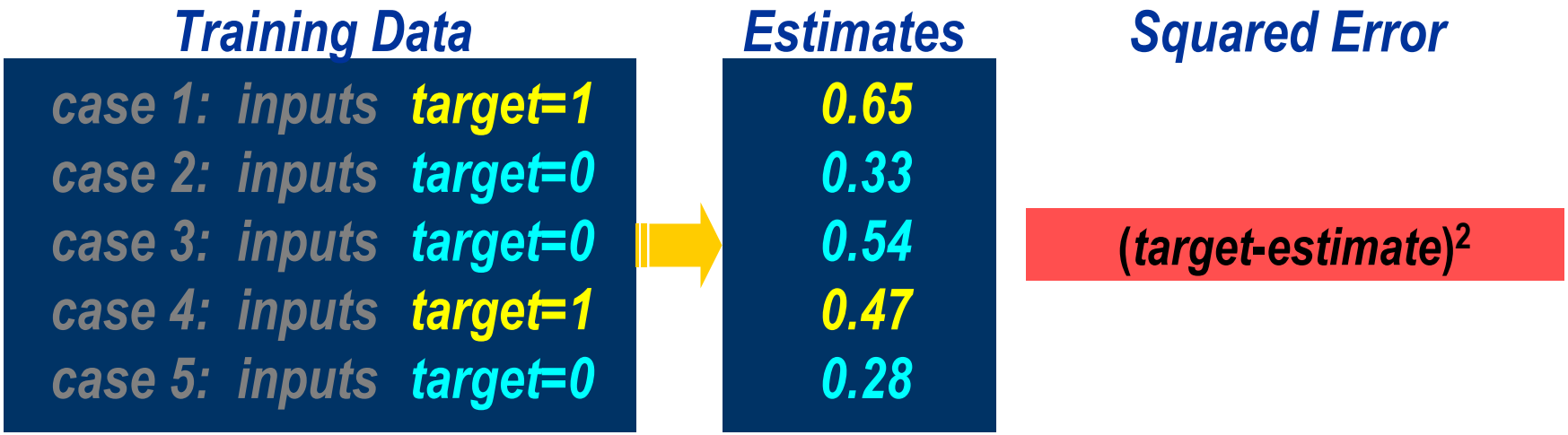
Ranking Assessment

Focus on correct ordering

Ranking Assessment (for Pessimists)

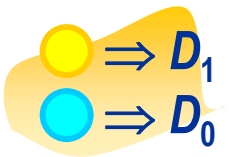
Focus on incorrect ordering

Estimate Assessment (only Pessimistic!)



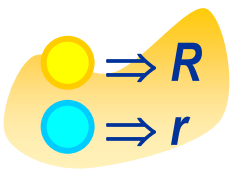
Focus on incorrect estimation

Predictive Modeling Assessments



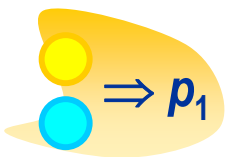
Decisions

Accuracy
Misclassification



Rankings

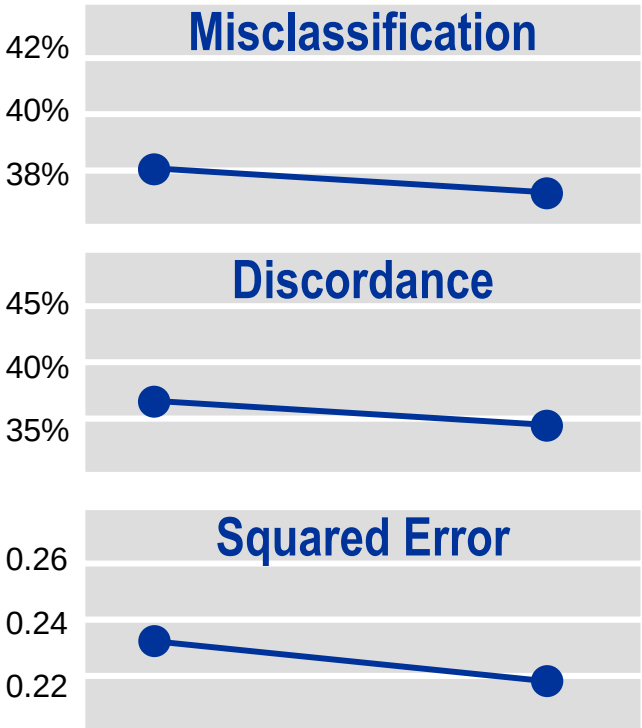
Concordance
Discordance



Estimates

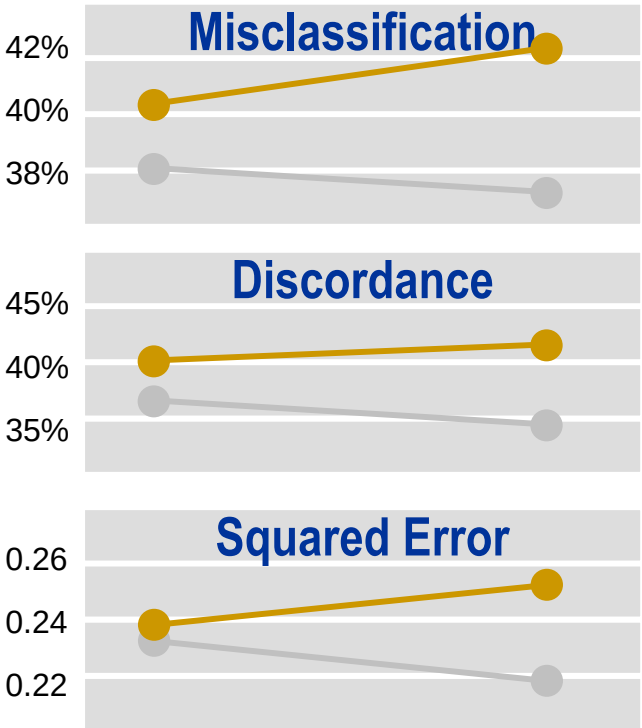
Squared Error

Optimistic Assessment, Pessimistic Stats



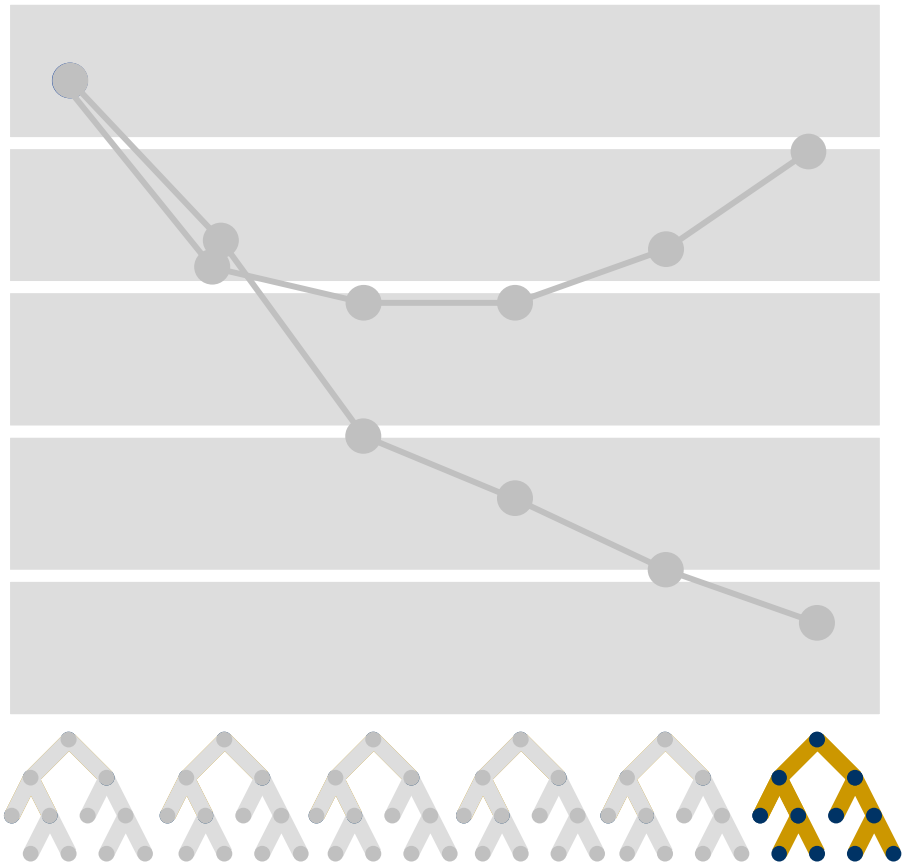
Increasing leaf count *improves* assessment measures on training data.

Unbiased Assessment



Increasing leaf count *might worsen* assessment measures on validation data.

Pruning Strategy

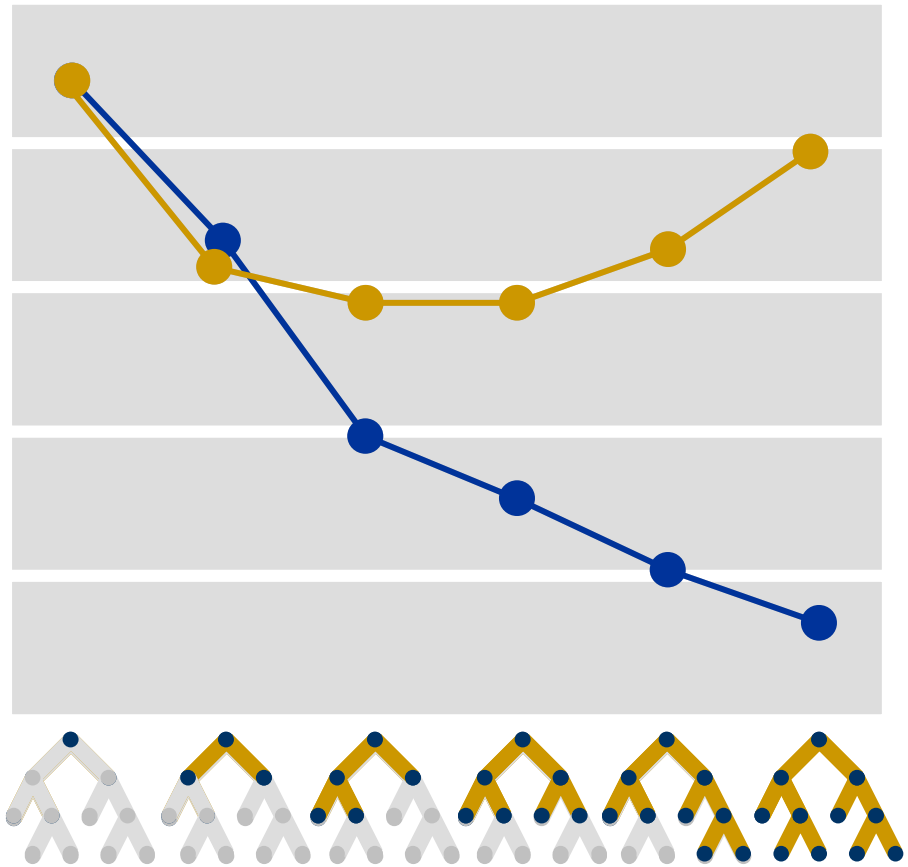


Create maximal tree.

Find optimal sub-trees.

Pick best sub-tree.

Pruning Strategy

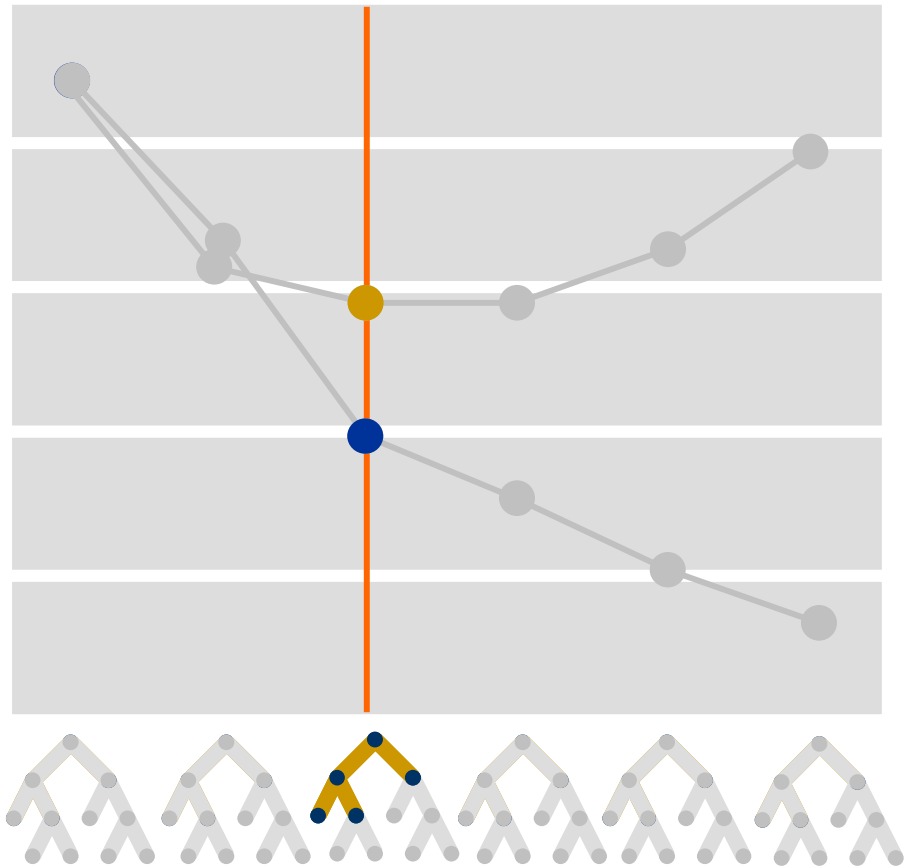


Create maximal tree.

Find optimal sub-trees.

Pick best sub-tree.

Pruning Strategy



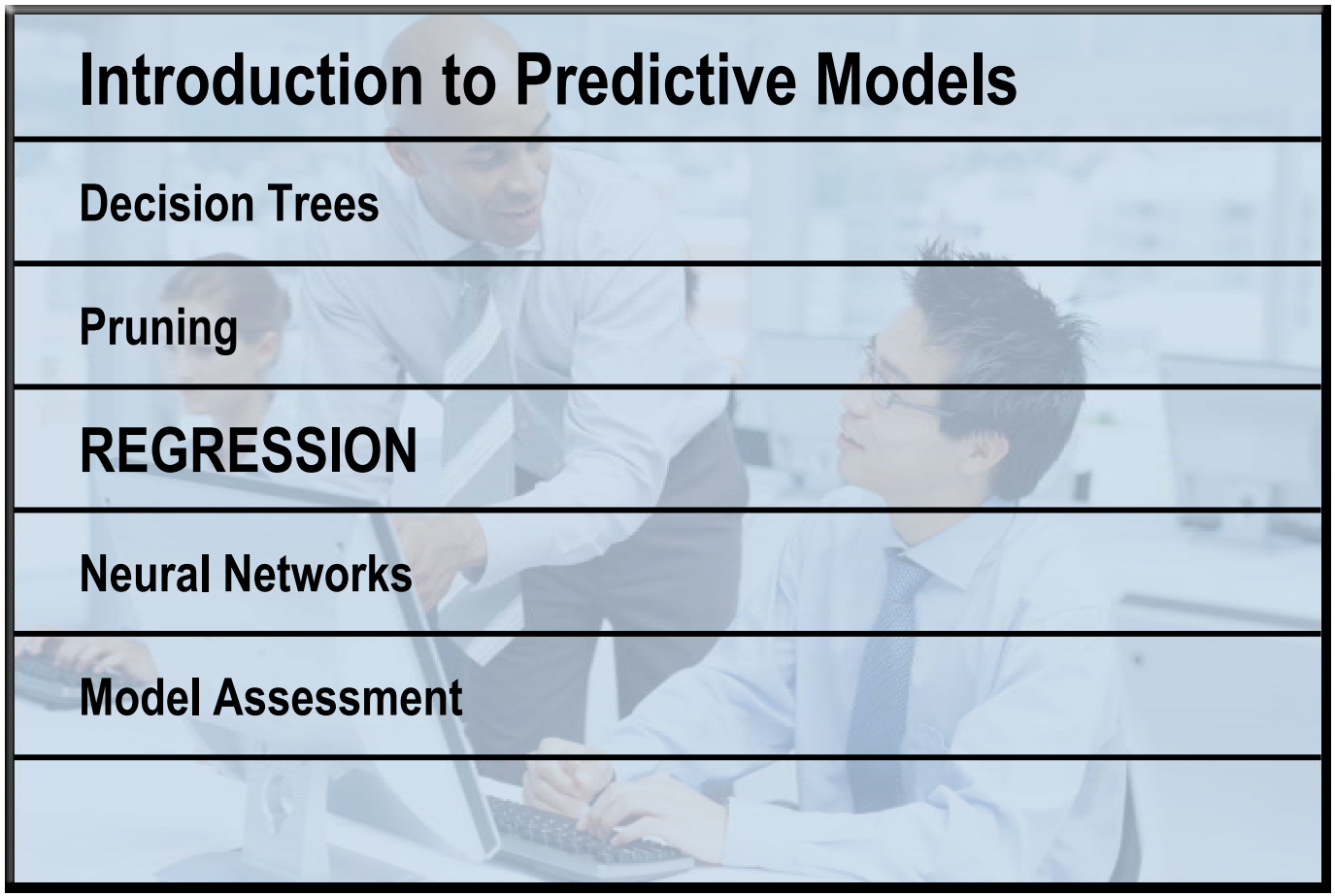
Create maximal tree.

Find optimal sub-trees.

Pick best sub-tree.

Demo Pruning

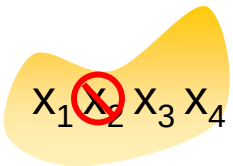
Agenda


Introduction to Predictive Models
Decision Trees
Pruning
REGRESSION
Neural Networks
Model Assessment

Model Essentials – Regressions

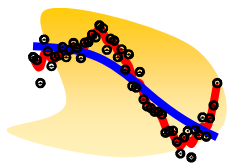
Predict new cases

Prediction formula



Select useful inputs

Sequential selection



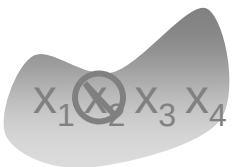
Optimize complexity

Optimal sequence model

Model Essentials – Regressions

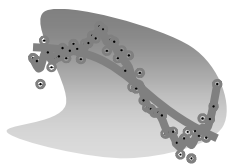
Predict new cases

Prediction formula



Select useful inputs

Sequential selection



Optimize complexity

Optimal sequence model

Linear Regression Prediction Formula

input measurement

$$\hat{y} = \hat{w}_0 + \hat{w}_1 \cdot x_1 + \hat{w}_2 \cdot x_2$$

prediction estimate

intercept estimate — $\hat{w}_0$

parameter estimate — $\hat{w}_1$, $\hat{w}_2$

Choose intercept and parameter estimates to *minimize*.

squared error function

$$\sum_{\text{training data}} (y_i - \hat{y}_i)^2$$

Logistic Regression Prediction Formula

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{w}_0 + \hat{w}_1 \cdot x_1 + \hat{w}_2 \cdot x_2$$

logit scores

Choose intercept and parameter estimates to *maximize*.

log-likelihood function

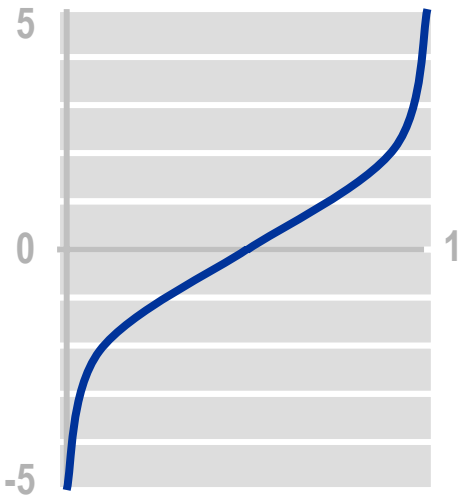
$$\sum_{\substack{\text{primary} \\ \text{outcome} \\ \text{training cases}}} \log(\hat{p}_i) + \sum_{\substack{\text{secondary} \\ \text{outcome} \\ \text{training cases}}} \log(1-\hat{p}_i)$$

Logit Link Function

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{w}_0 + \hat{w}_1 \cdot x_1 + \hat{w}_2 \cdot x_2$$

logit scores

logit link function



Δx_i	consequence
1	$\Rightarrow$ odds $\times$ $\exp(w_i)$
$\frac{0.69}{w_i}$	$\Rightarrow$ odds $\times 2$

odds ratio

doubling amount

Model interpretation

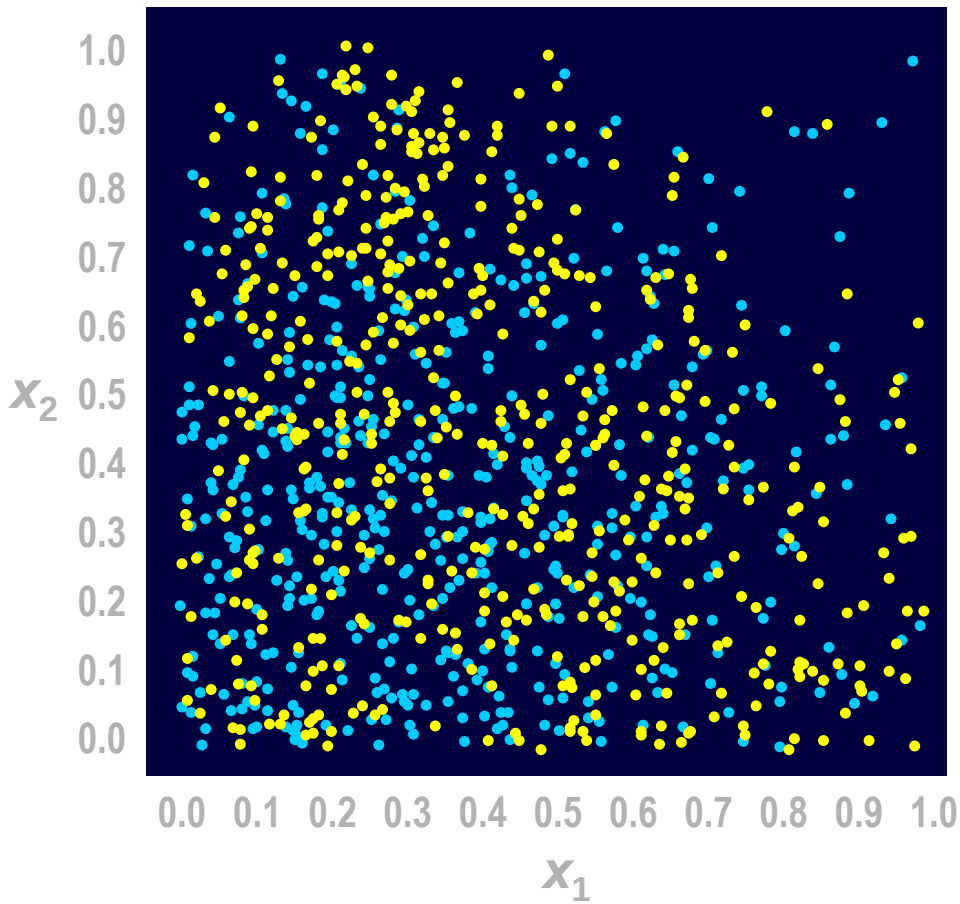
Simple Prediction Illustration – Regressions

logit equation

$$\text{logit}(\hat{p}) = \hat{w}_0 + \hat{w}_1 \cdot x_1 + \hat{w}_2 \cdot x_2$$

$$\hat{p} = \frac{1}{1 + e^{-\text{logit}(\hat{p})}}$$

logistic equation

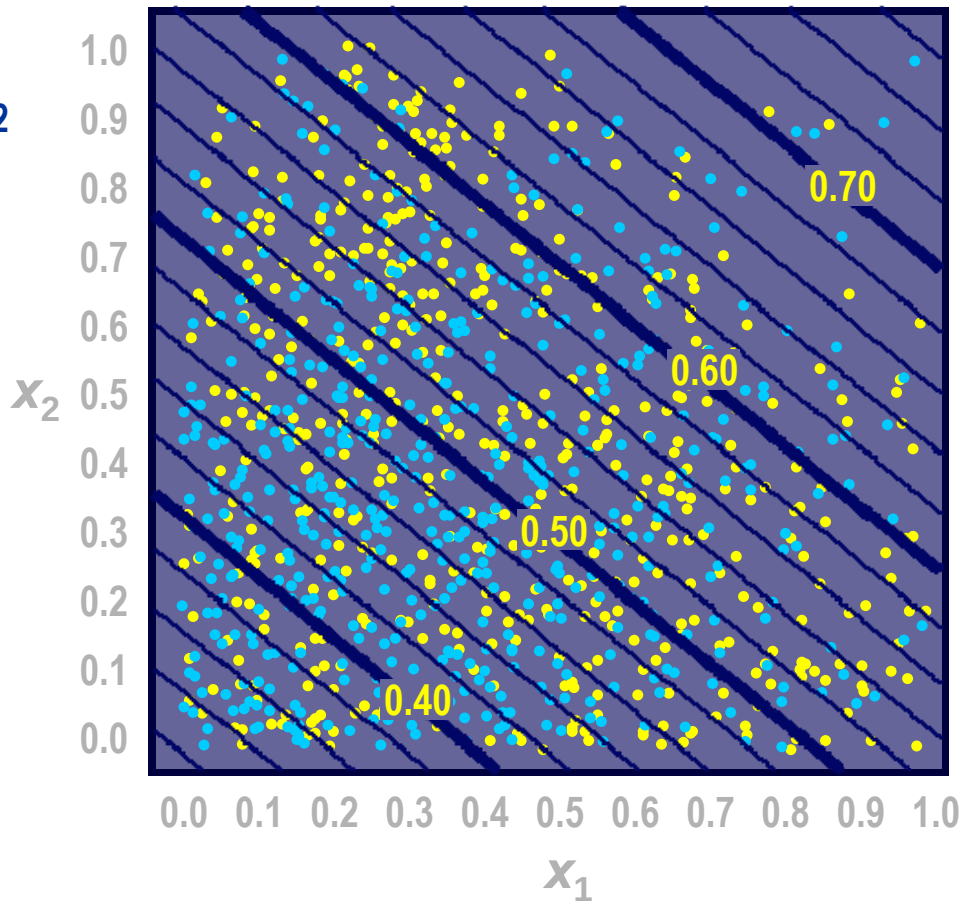


Prediction Estimates

$$\text{logit}(\hat{p}) = -0.81 + 0.92x_1 + 1.11x_2$$

$$\hat{p} = \frac{1}{1 + e^{0.81 - 0.92 \cdot x_1 - 1.11 \cdot x_2}}$$

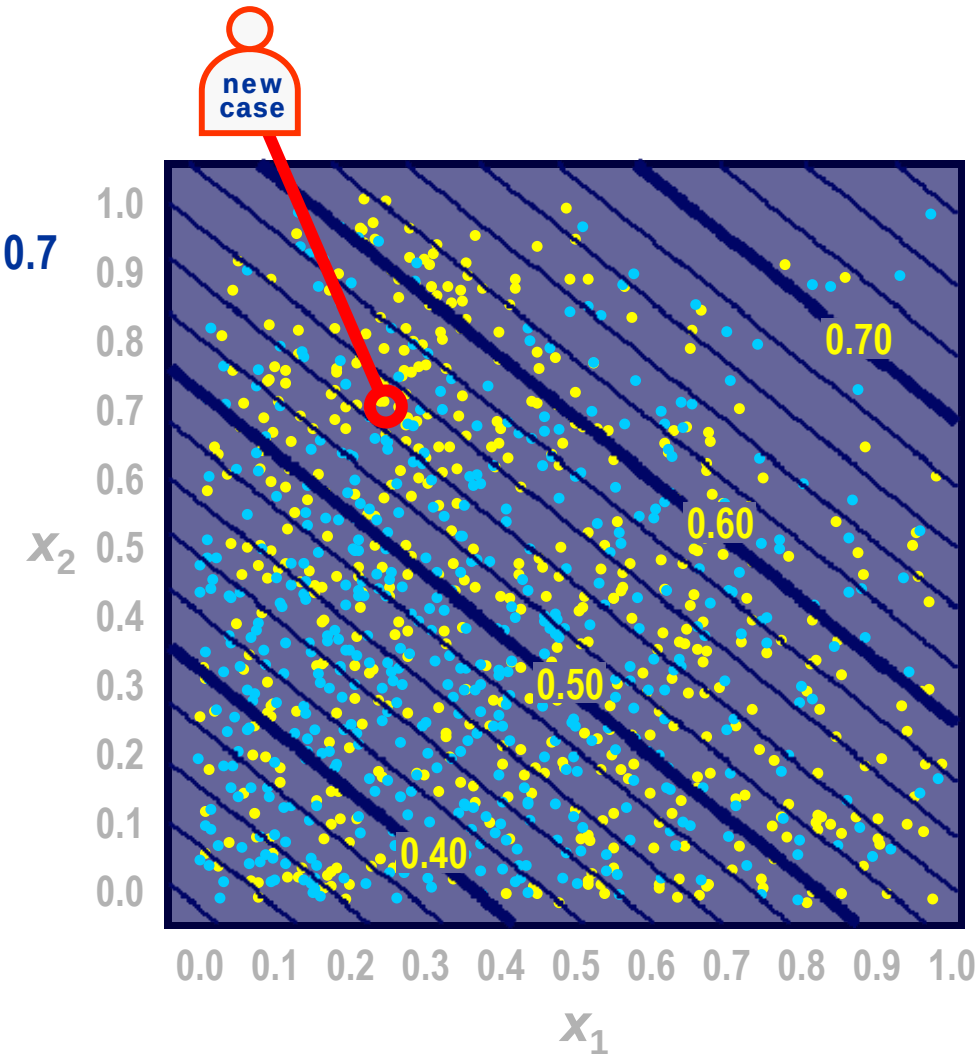
	x_1	x_2
odds ratio	2.5	3.0
doubling amount	0.75	0.62



Prediction Estimates

$$\text{logit}(\hat{p}) = -0.81 + 0.92\ 0.3 + 1.11\ 0.7$$

$$\hat{p} = \frac{1}{1 + e^{\,0.81 - 0.92\ 0.3 - 1.11\ 0.7}}$$



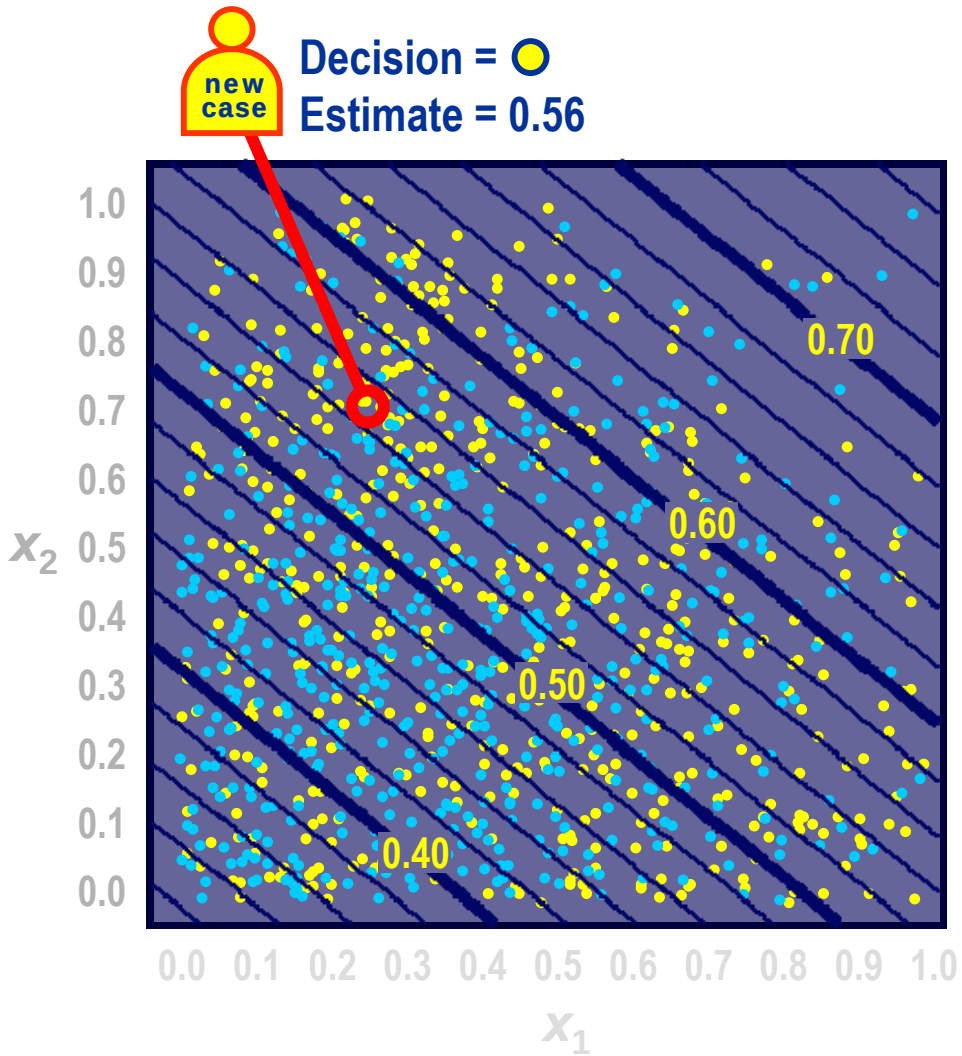
Prediction Estimates

logit score

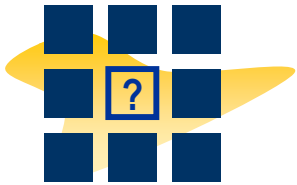
$\text{logit}(\hat{p}) = 0.104$

$\hat{p} = 0.56$

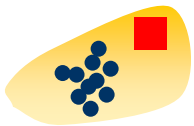
prediction estimate



Beyond the Prediction Formula

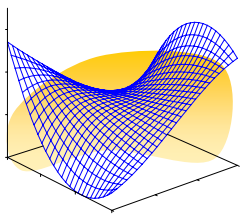


Missing values



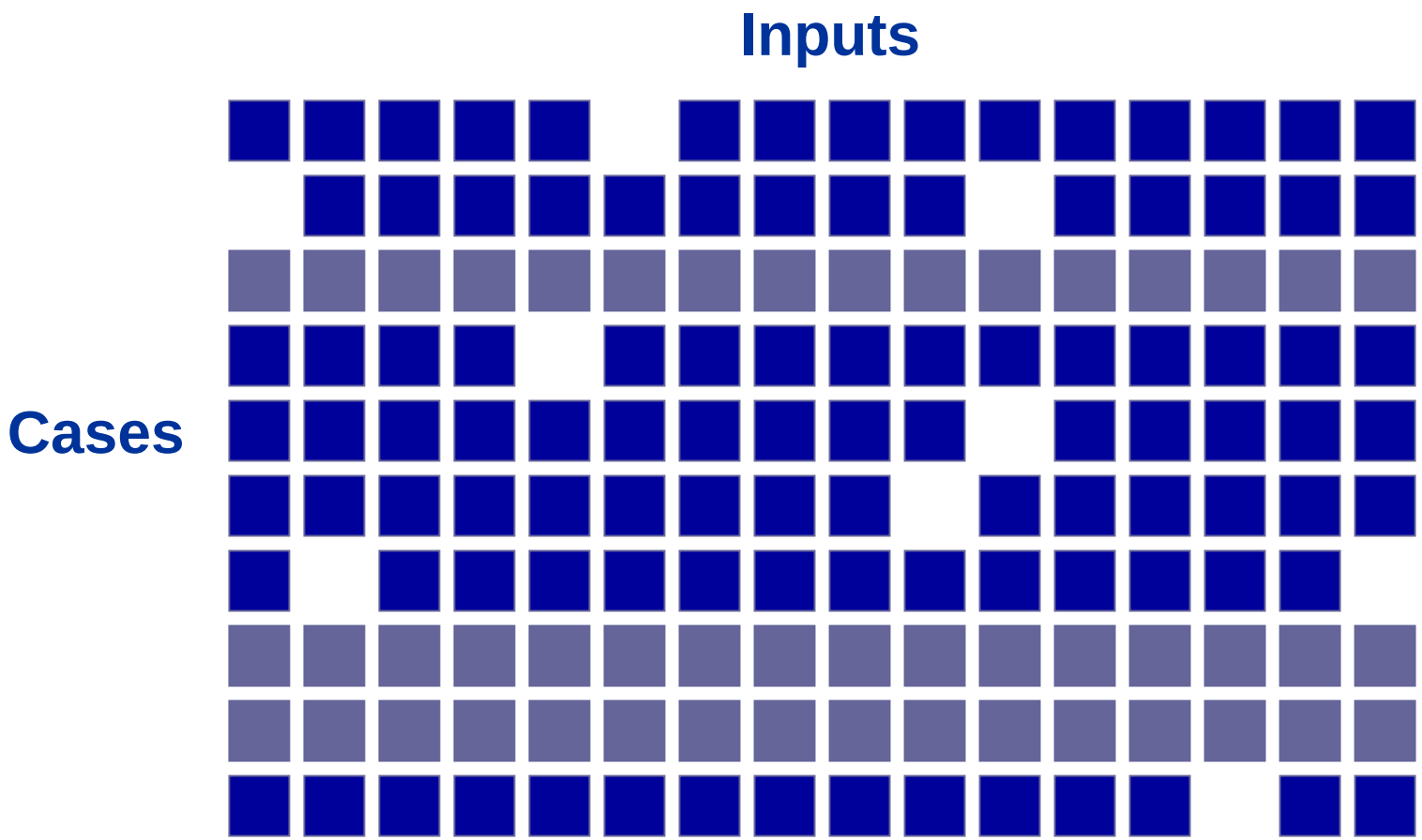
Extreme or unusual values

Non-numeric inputs

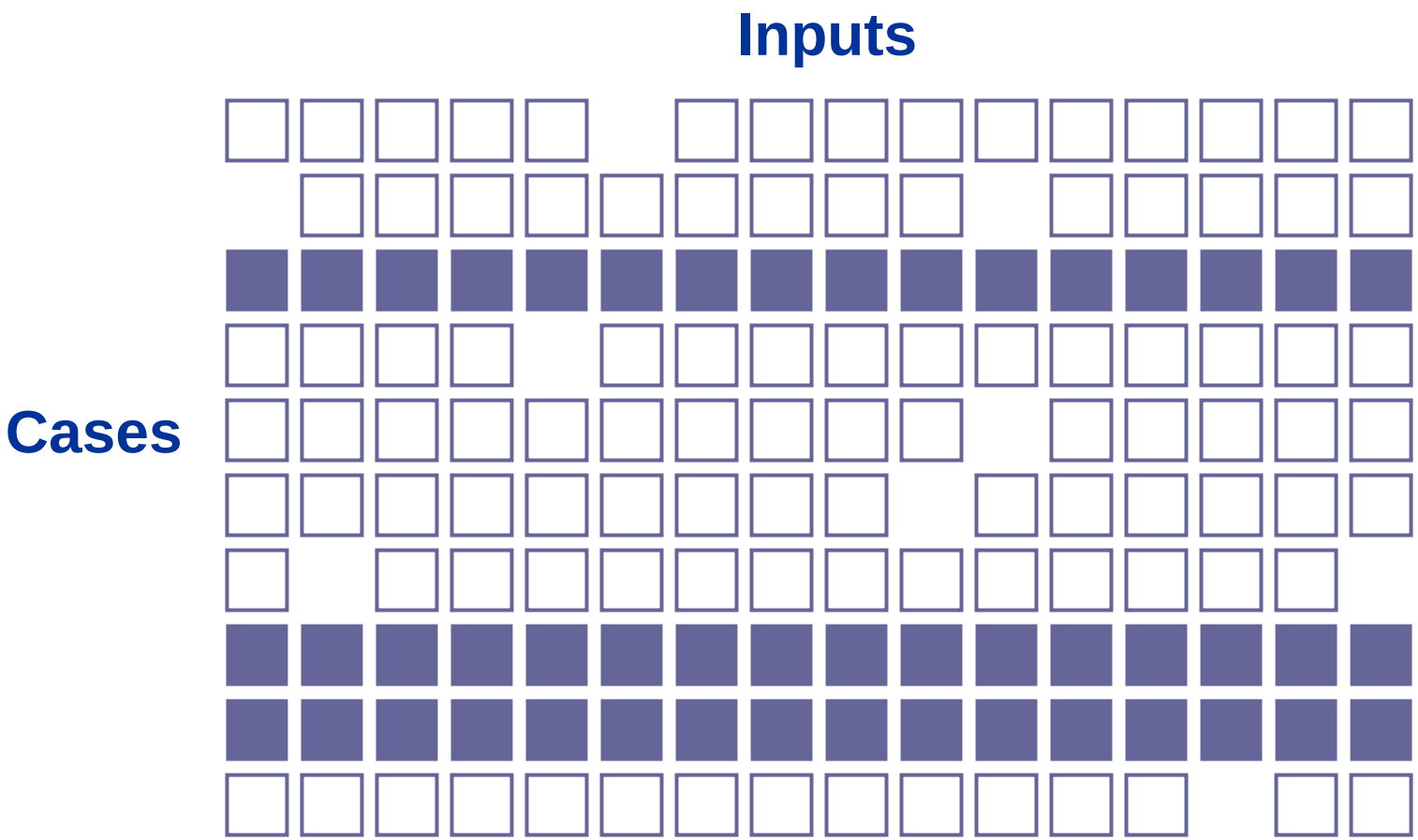


Nonlinearity and Non-additivity

Missing Values and Regression Modeling



Missing Values and Regression Modeling



Missing Values and the Prediction Formula

Prediction Formula:

$$\text{logit}(p) = -2.1 + 0.072x_1 - 0.89x_2 - 1.24x_3$$

New Case:

$$(x_1, x_2, x_3) = (2, ?, -1)$$

Predicted Value:

$$\text{logit}(p) = -2.1 + 0.144x_1 - 0.8\,? + 1.24$$

Missing Value Causes

Not applicable

No match

Non-disclosure

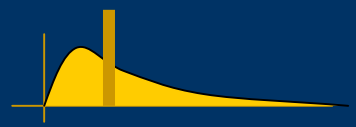
Missing Value Remedies

Not applicable

No match

Non-disclosure

Synthetic distribution



Estimation

$x_i = f(x_1, \ldots, x_p)$

Managing Missing Values

This demonstration illustrates how to impute synthetic data values and create missing value indicators.

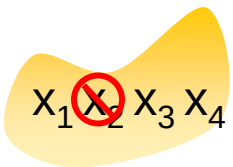
Running the Regression Node

This demonstration illustrates using the Regression tool.

Model Essentials – Regressions

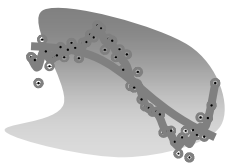
Predict new cases

Prediction formula



Select useful inputs

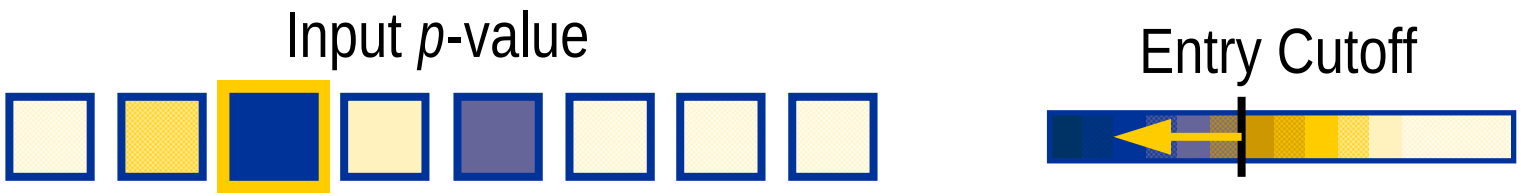
Sequential selection



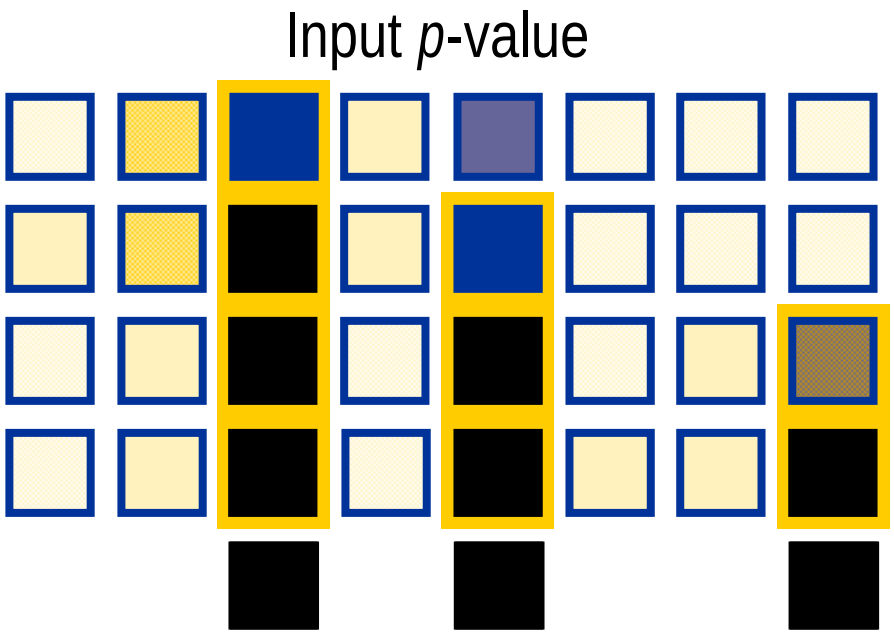
Optimize complexity

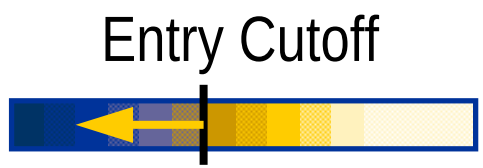
Optimal sequence model

Sequential Selection – Forward

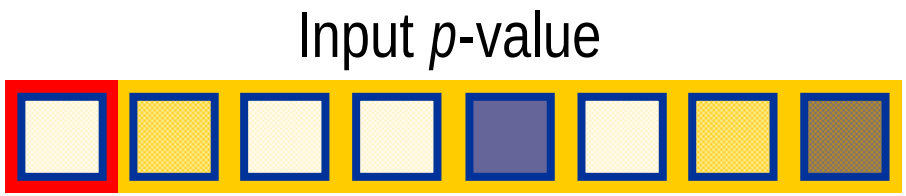


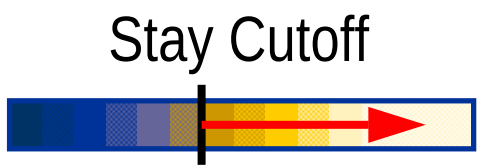
Sequential Selection – Forward



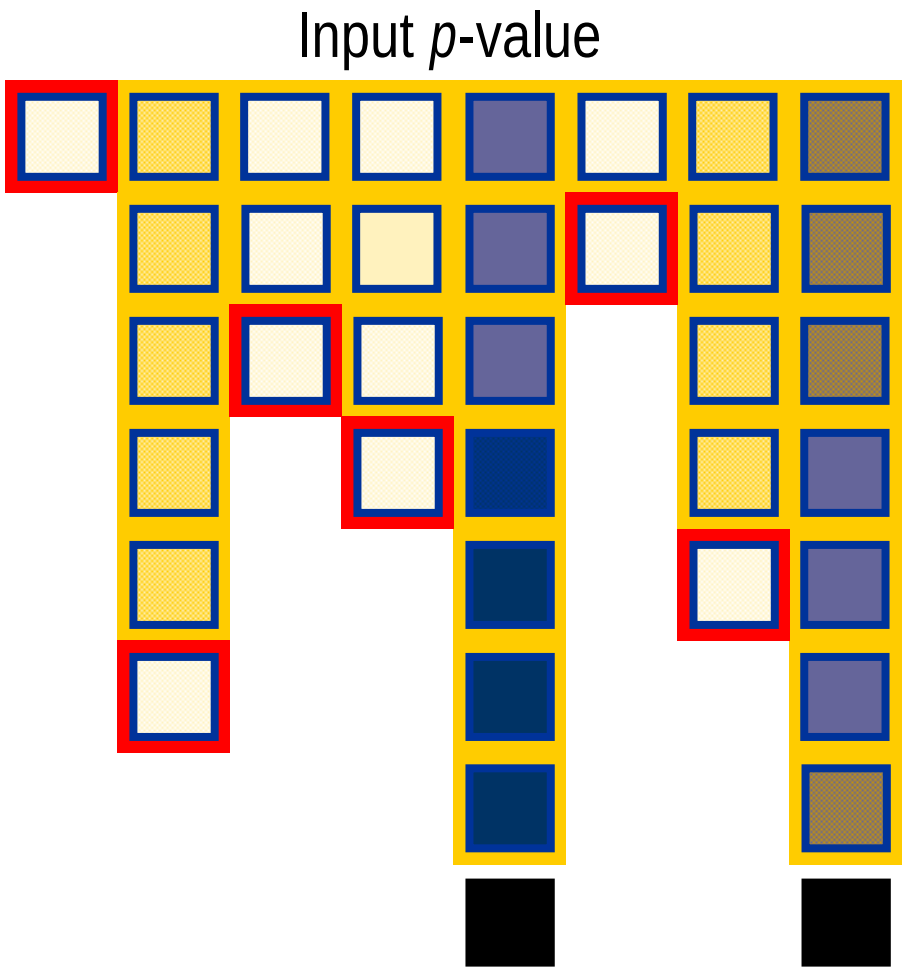


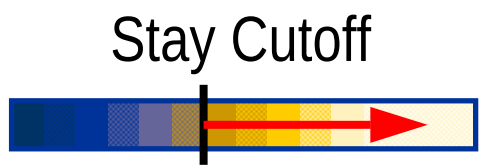
Sequential Selection – Backward



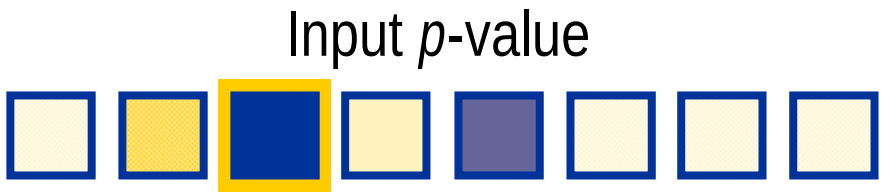


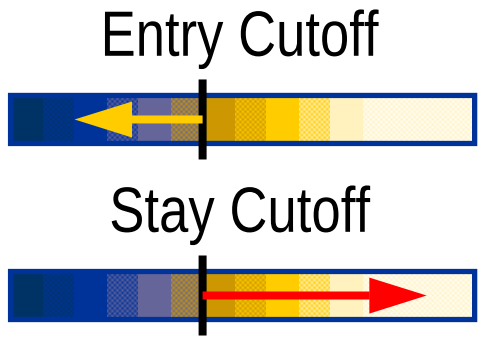
Sequential Selection – Backward



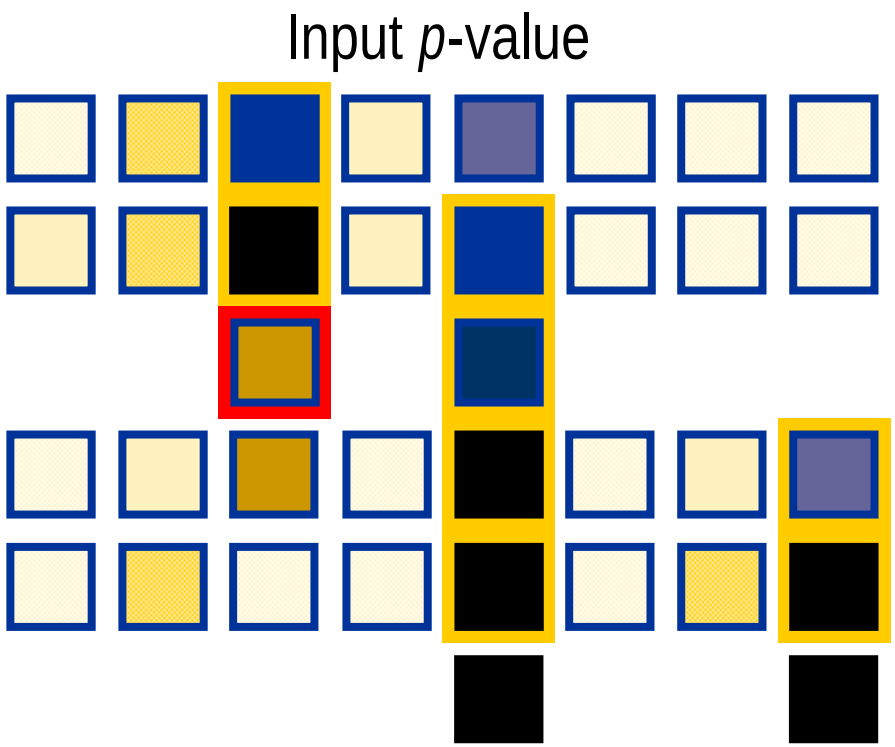


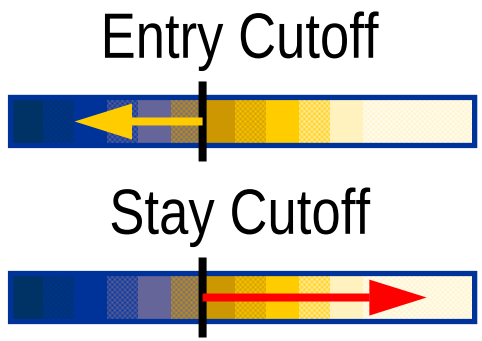
Sequential Selection – Stepwise





Sequential Selection – Stepwise

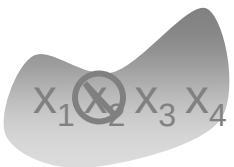




Model Essentials – Regressions

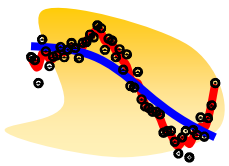
Predict new cases

Prediction formula



Select useful inputs

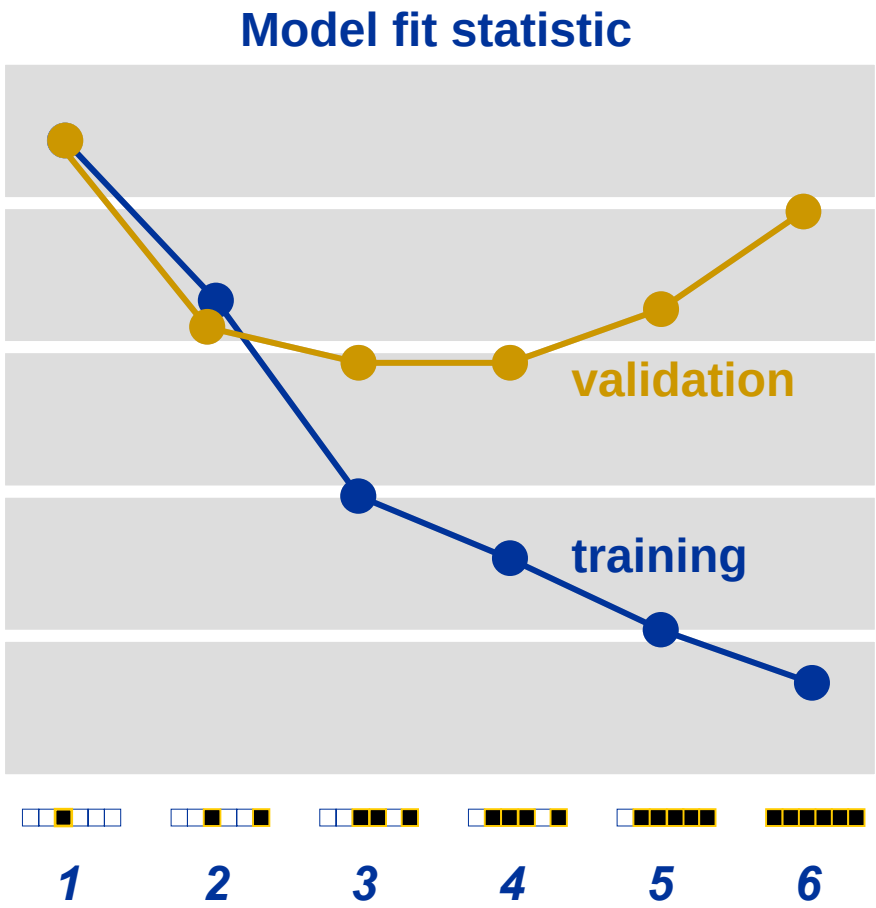
Sequential selection



Optimize complexity

Optimal sequence model

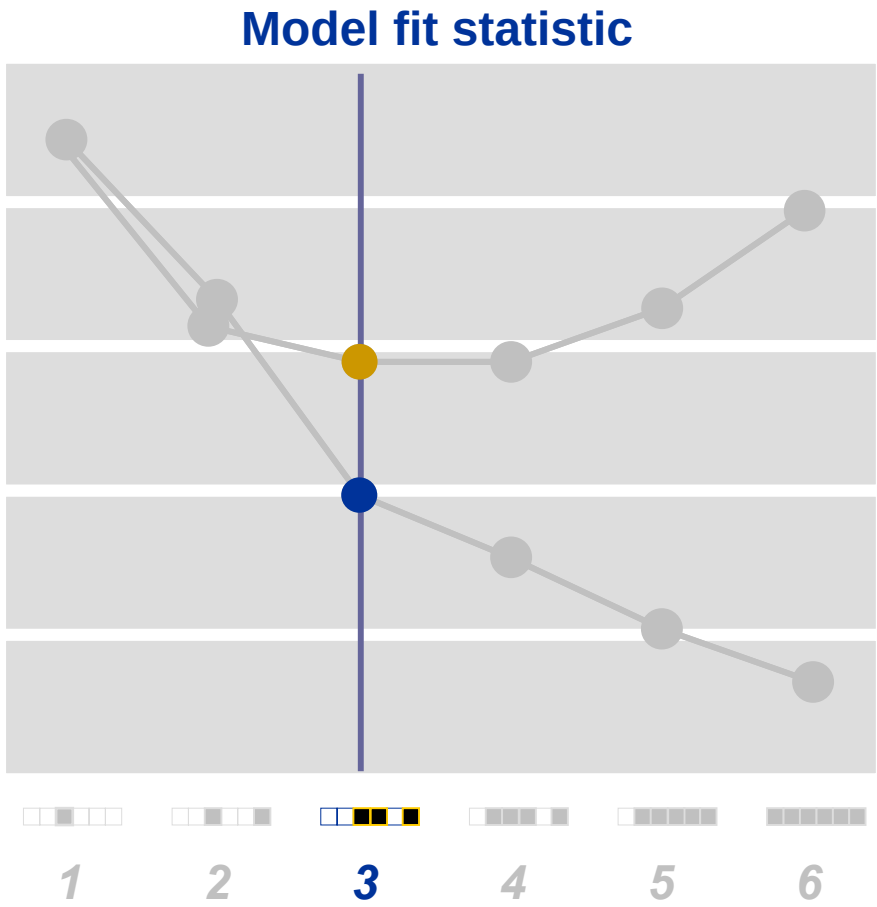
Model Fit versus Complexity



Evaluate each sequence step.

Choose simplest optimal model.

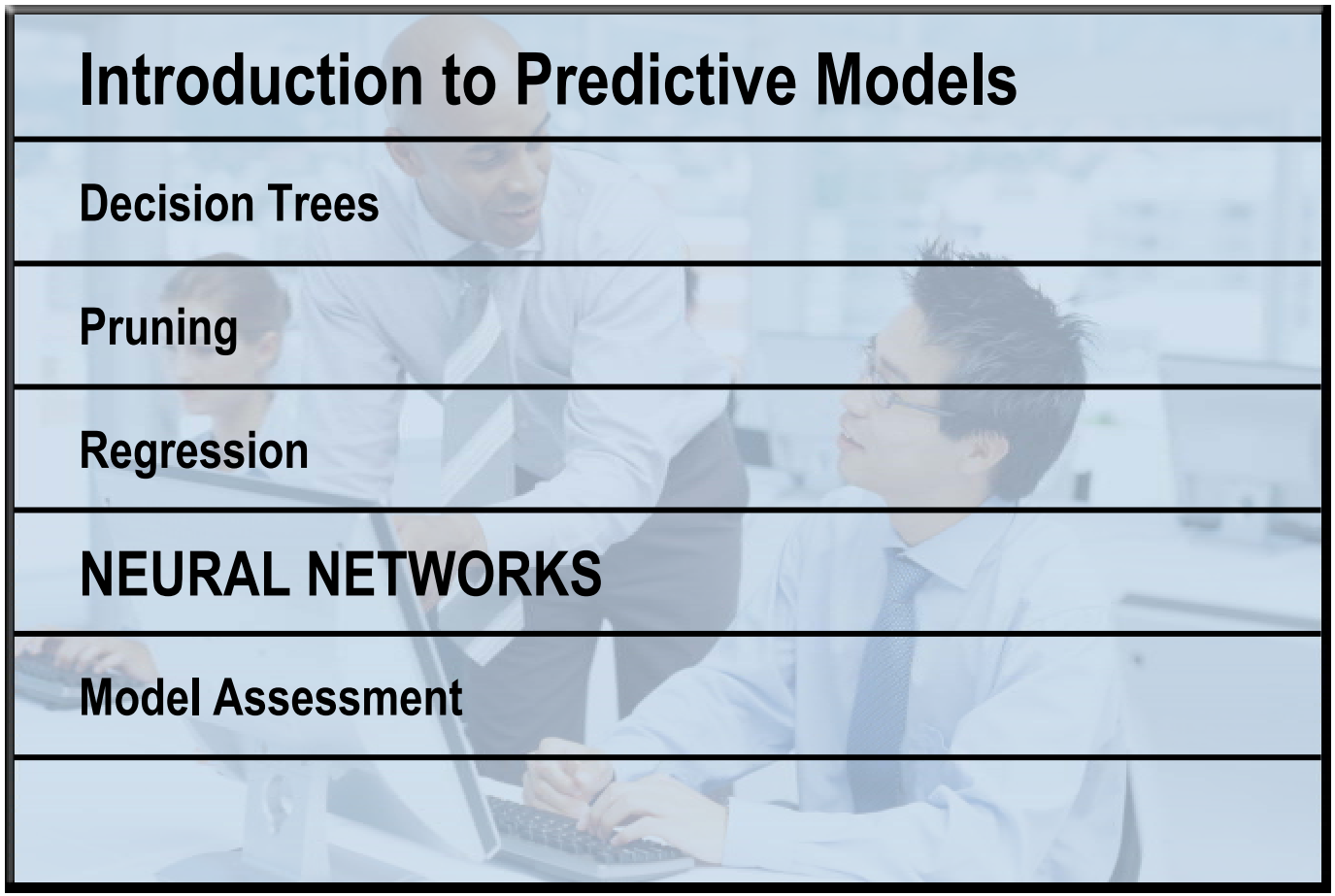
Select Model with Optimal Validation Fit



Evaluate each sequence step.

Choose simplest optimal model.

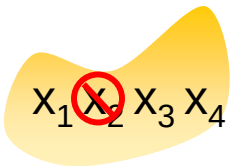
Agenda


Introduction to Predictive Models
Decision Trees
Pruning
Regression
NEURAL NETWORKS
Model Assessment

Model Essentials – Neural Networks

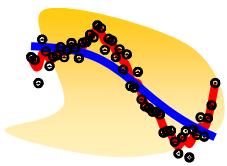
Predict new cases

Prediction formula



Select useful inputs

None



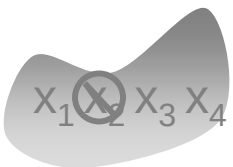
Optimize complexity

Stopped training

Model Essentials – Neural Networks

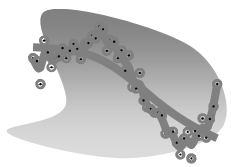
Predict new cases

Prediction formula



Select useful inputs

None



Optimize complexity

Stopped training

Neural Network Prediction Formula

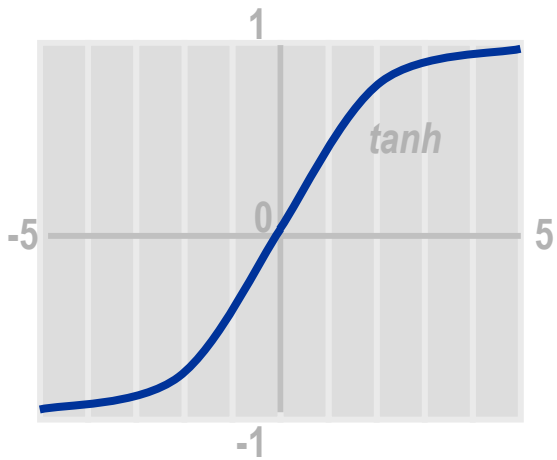
hidden unit

prediction estimate

$$\hat{y} = \hat{w}_{00} + \hat{w}_{01} \cdot H_1 + \hat{w}_{02} \cdot H_2 + \hat{w}_{03} \cdot H_3$$

bias estimate

weight estimate



$$H_1 = \tanh(\hat{w}_{10} + \hat{w}_{11} x_1 + \hat{w}_{12} x_2)$$

$$H_2 = \tanh(\hat{w}_{20} + \hat{w}_{21} x_1 + \hat{w}_{22} x_2)$$

$$H_3 = \tanh(\hat{w}_{30} + \hat{w}_{31} x_1 + \hat{w}_{32} x_2)$$

activation function

Neural Network Binary Prediction Formula

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{w}_{00} + \hat{w}_{01} H_1 + \hat{w}_{02} H_2 + \hat{w}_{03} H_3$$

logit link function

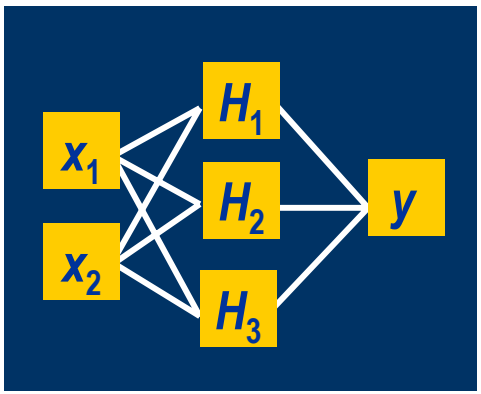
$$H_1 = \tanh(\hat{w}_{10} + \hat{w}_{11} x_1 + \hat{w}_{12} x_2)$$

$$H_2 = \tanh(\hat{w}_{20} + \hat{w}_{21} x_1 + \hat{w}_{22} x_2)$$

$$H_3 = \tanh(\hat{w}_{30} + \hat{w}_{31} x_1 + \hat{w}_{32} x_2)$$

Neural Network Diagram

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{w}_{00} + \hat{w}_{01} H_1 + \hat{w}_{02} H_2 + \hat{w}_{03} H_3$$



$$H_1 = \tanh(\hat{w}_{10} + \hat{w}_{11} x_1 + \hat{w}_{12} x_2)$$

$$H_2 = \tanh(\hat{w}_{20} + \hat{w}_{21} x_1 + \hat{w}_{22} x_2)$$

$$H_3 = \tanh(\hat{w}_{30} + \hat{w}_{31} x_1 + \hat{w}_{32} x_2)$$

Prediction Illustration – Neural Networks

logit equation

$$\text{logit}(\hat{p}) = \hat{w}_{00} + \hat{w}_{01}H_1 + \hat{w}_{02}H_2 + \hat{w}_{03}H_3$$

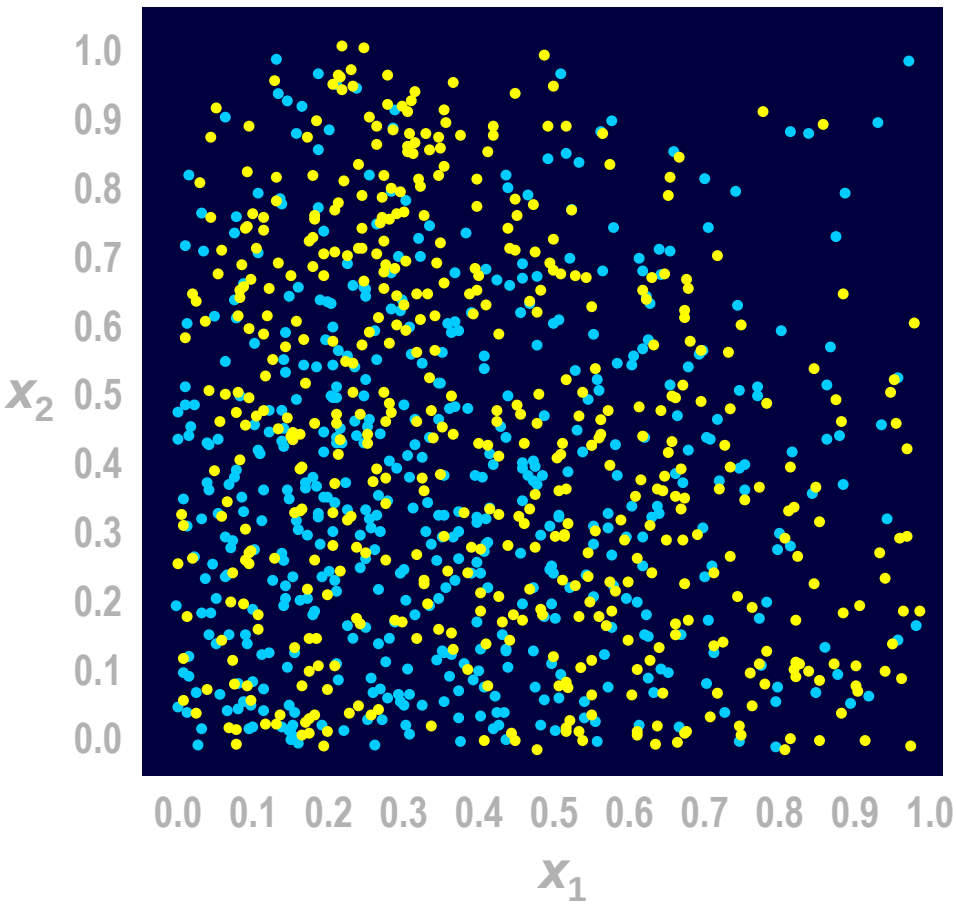
$$H_1 = \tanh(\hat{w}_{10} + \hat{w}_{11}x_1 + \hat{w}_{12}x_2)$$

$$H_2 = \tanh(\hat{w}_{20} + \hat{w}_{21}x_1 + \hat{w}_{22}x_2)$$

$$H_3 = \tanh(\hat{w}_{30} + \hat{w}_{31}x_1 + \hat{w}_{32}x_2)$$

$$\hat{p} = \frac{1}{1 + e^{-\text{logit}(\hat{p})}}$$

logistic equation



Prediction Estimates

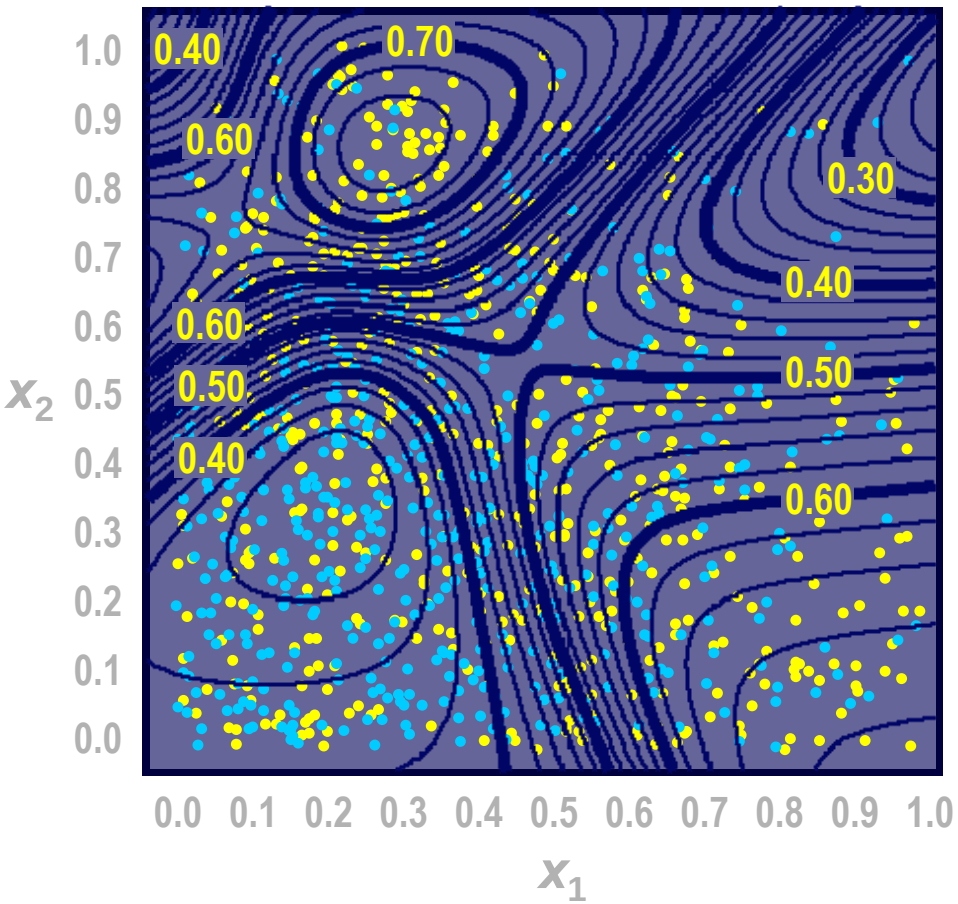
$$\text{logit}(\hat{p}) = -0.5 - 2.6H_1 - 1.9H_2 + .63H_3$$

$$H_1 = \tanh(-1.8 - .25x_1 + 2.4x_2)$$

$$H_2 = \tanh(2.7 + 2.7x_1 - 5.3x_2)$$

$$H_3 = \tanh(-5.0 + 8.1x_1 + 4.3x_2)$$

$$\hat{p} = \frac{1}{1 + e^{-\text{logit}(\hat{p})}}$$



Prediction Estimates

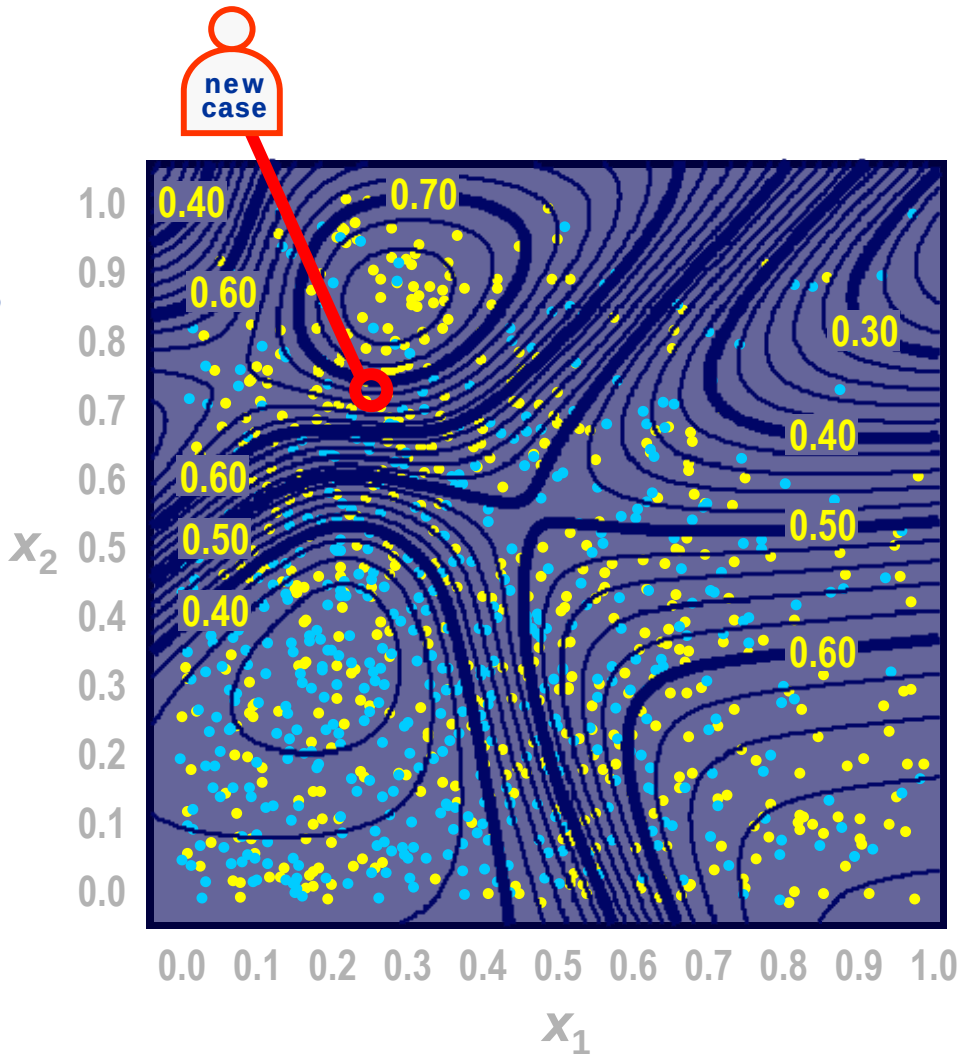
$$\text{logit}(\hat{p}) = -0.5 - 2.6H_1 - 1.9H_2 + .63\,H_3$$

$H_1 = -0.209$

$H_2 = -0.231$

$H_3 = 0.457$

$$\hat{p} = \frac{1}{1 + e^{-\text{logit}(\hat{p})}}$$



Prediction Estimates

logit score

$\text{logit}(\hat{p}) = 0.774$

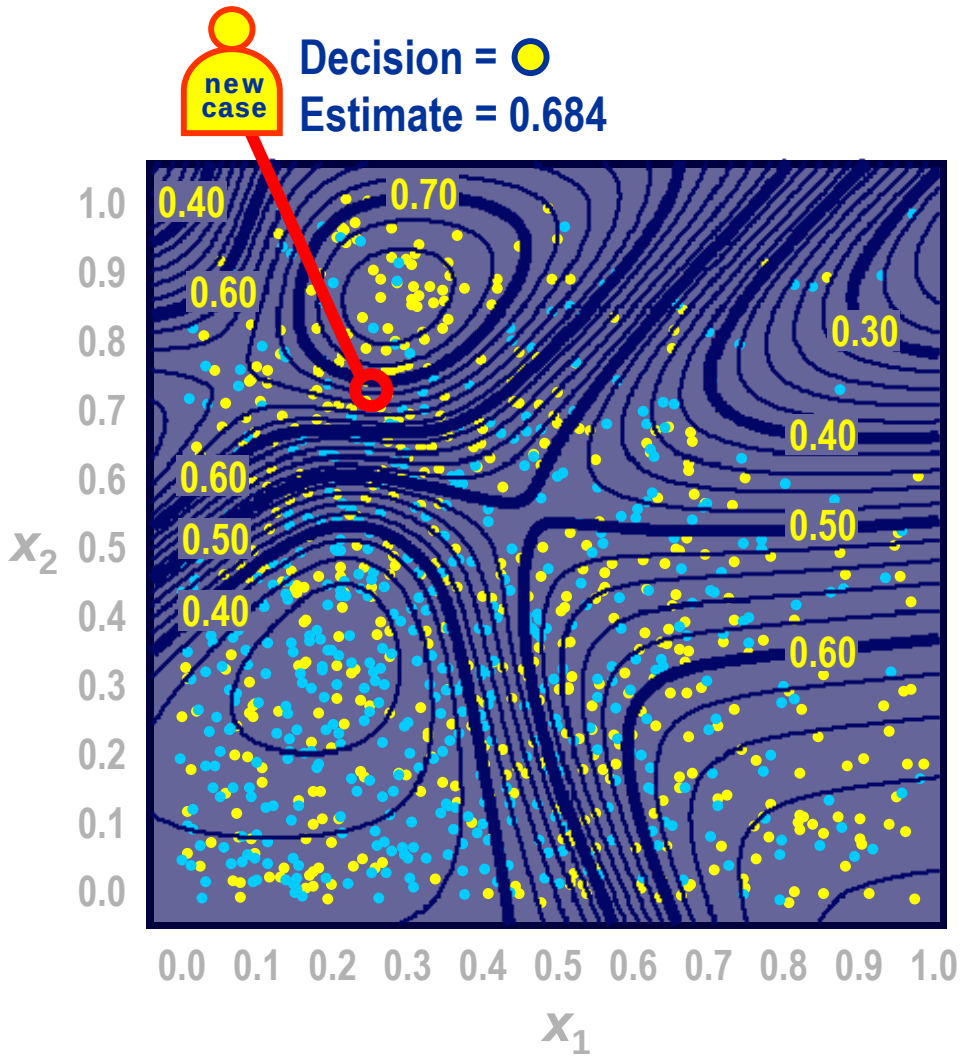
$H_1 = -0.209$

$H_2 = -0.231$

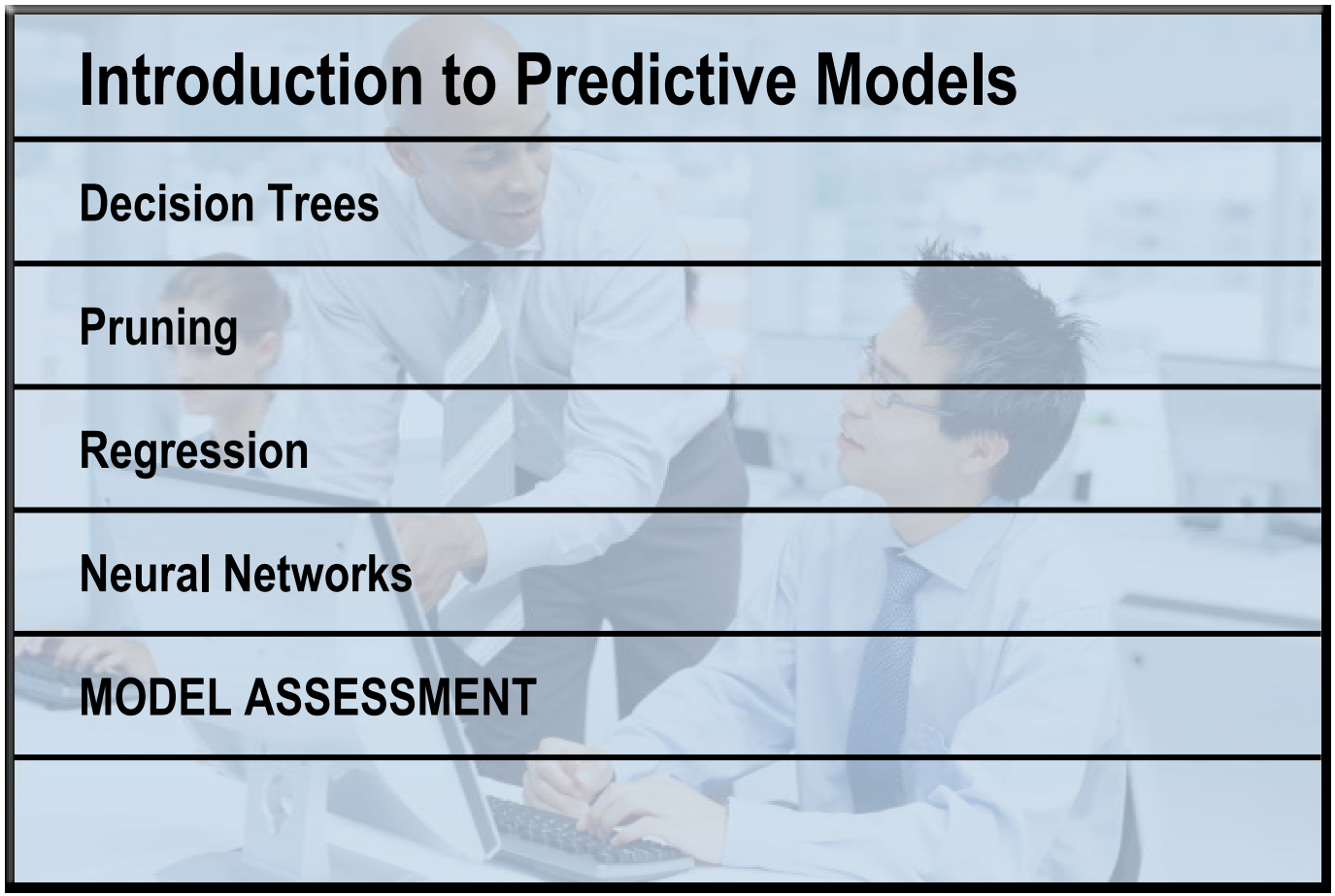
$H_3 = 0.457$

$\hat{p} = 0.684$

prediction estimate



Agenda


Introduction to Predictive Models
Decision Trees
Pruning
Regression
Neural Networks
MODEL ASSESSMENT

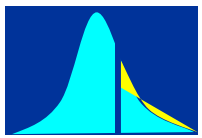
Assessment Types

The Model Comparison tool provides

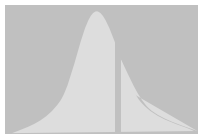
Summary statistics

Statistical graphics

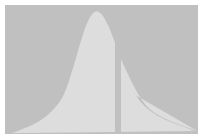
Summary Statistics Summary

Prediction Type	*Statistic*
 **Decisions**	**Accuracy / Misclassification** **Profit / Loss** **KS-statistic**
1,2,3,... **Rankings**	ROC Index (concordance) Gini coefficient
$\hat{p} \approx E(Y)$ **Estimates**	Average squared error SBC / Likelihood

Summary Statistics Summary

Prediction Type	*Statistic*
 Decisions	Accuracy / Misclassification Profit / Loss KS-statistic
1,2,3,… **Rankings**	**ROC Index (concordance)** **Gini coefficient**
$\hat{p} \approx E(Y)$ Estimates	Average squared error SBC / Likelihood

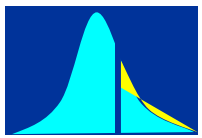
Summary Statistics Summary

Prediction Type	*Statistic*
 Decisions	Accuracy / Misclassification Profit / Loss KS-statistic
1,2,3,... Rankings	ROC Index (concordance) Gini coefficient
$\hat{p} \approx E(Y)$ **Estimates**	**Average squared error** **SBC / Likelihood**

Comparing Models with Summary Statistics

This demonstration illustrates the use of the Model Comparison tool, which collects assessment information from attached modeling nodes and enables you to easily compare model performance measures.

Statistical Graphics Summary

Prediction Type	*Statistic*
 Decisions	
1,2,3,… Rankings	Sensitivity charts Response rate charts
$\hat{p} \approx E(Y)$ Estimates	