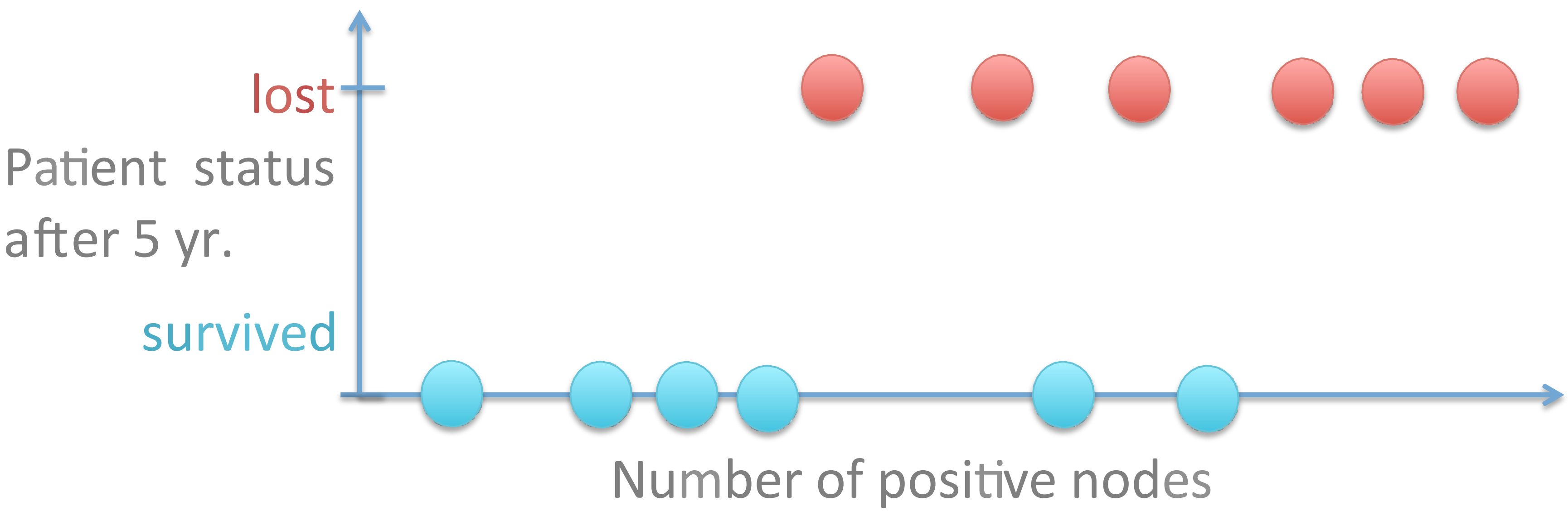
Logistic Regression

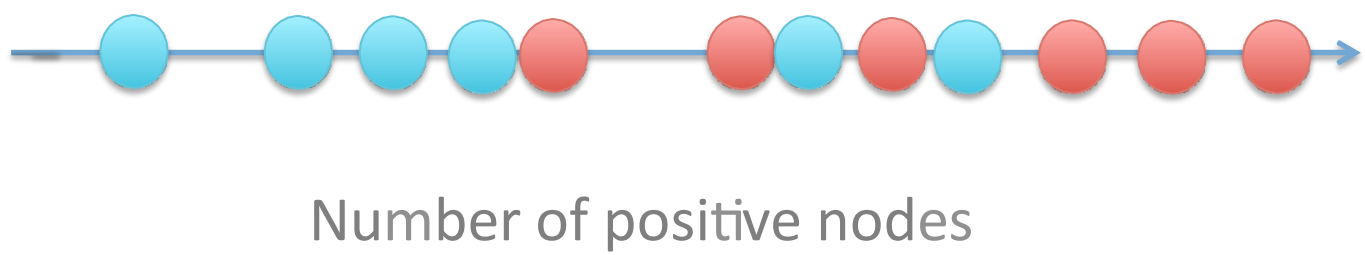
1 Feature: Number of + nodes

2 Labels: Survived / Lost



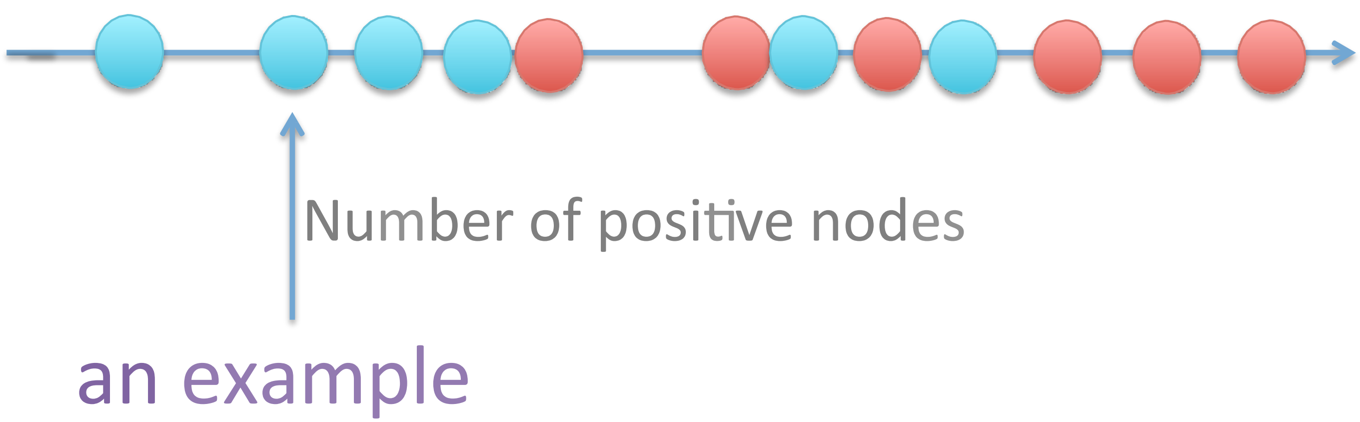
1 Feature: Number of + nodes

2 Labels: Survived / Lost



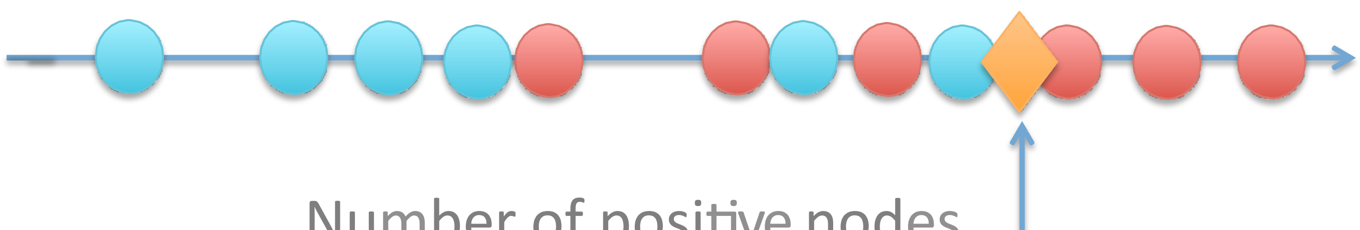
1 Feature: Number of + nodes

2 Labels: Survived / Lost



1 Feature: Number of + nodes

2 Labels: Survived / Lost

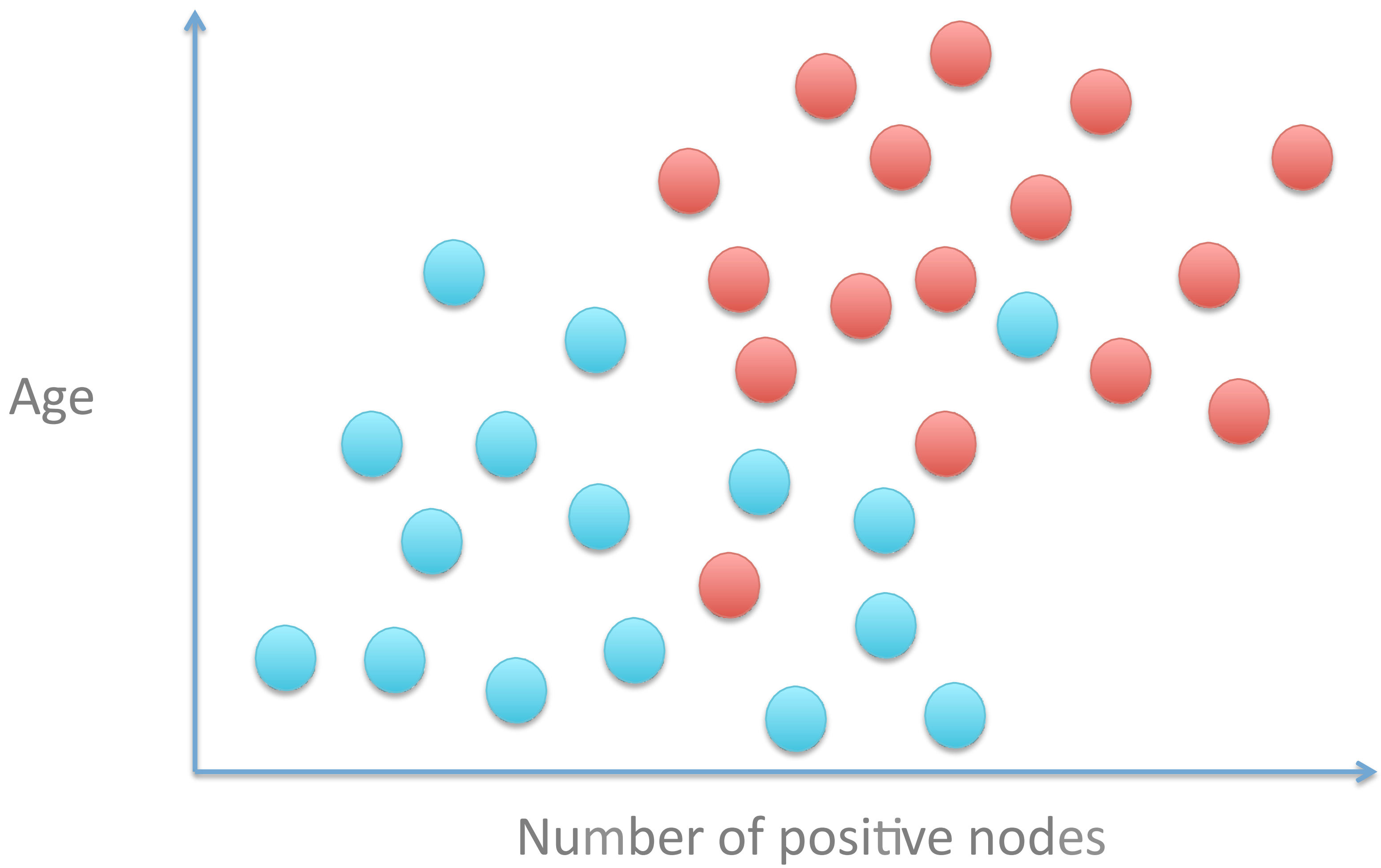


Number of positive nodes

new example (predict)

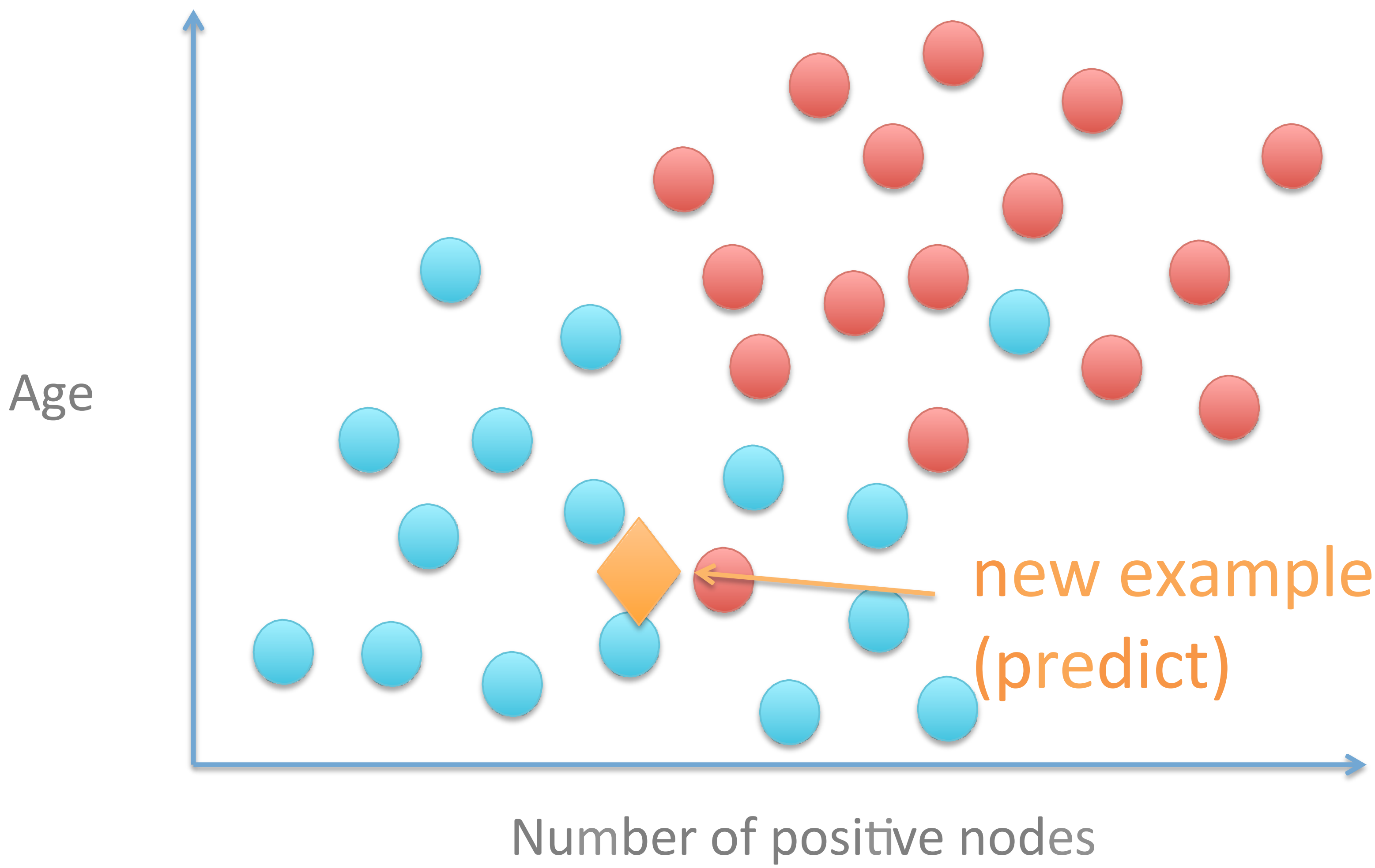
2 Features: Number of + nodes, Age

2 Labels: Survived / Lost

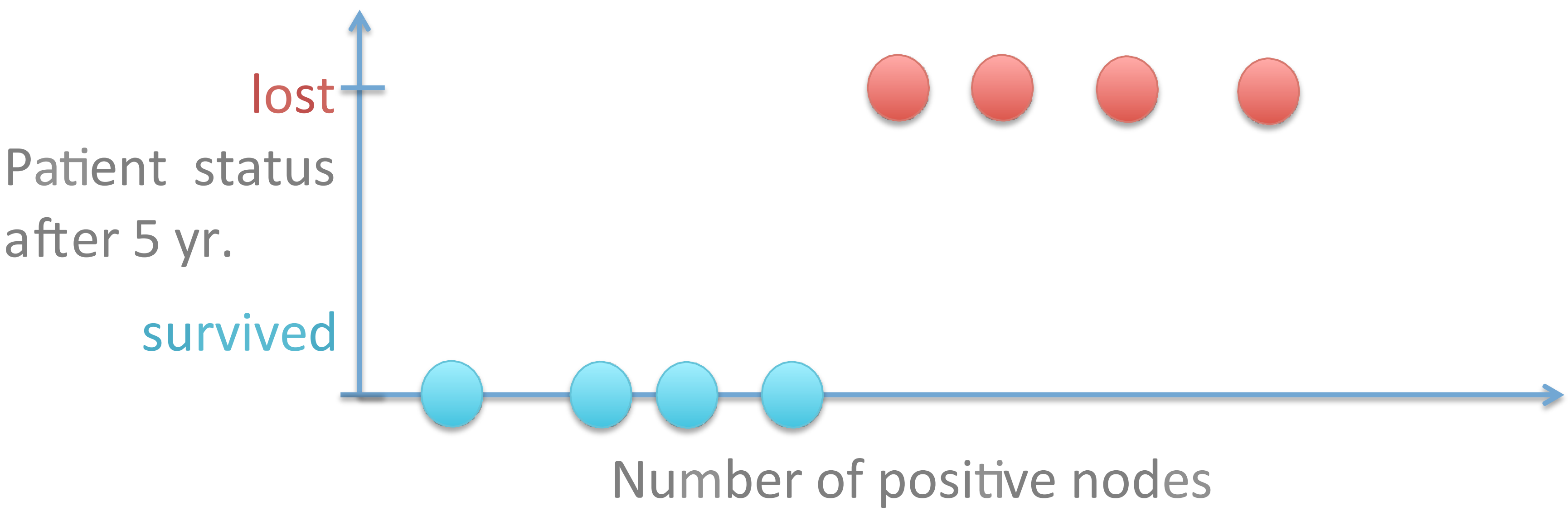


2 Features: Number of + nodes, Age

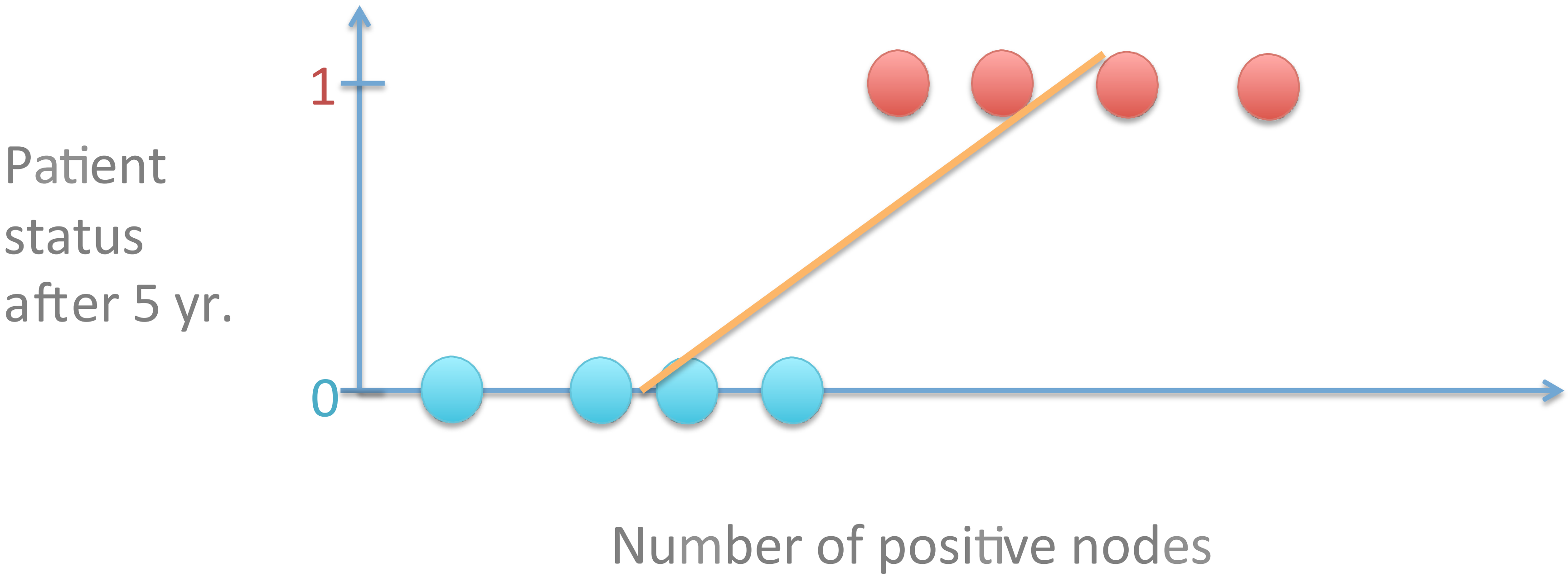
2 Labels: Survived / Lost



Linear regression for classification?

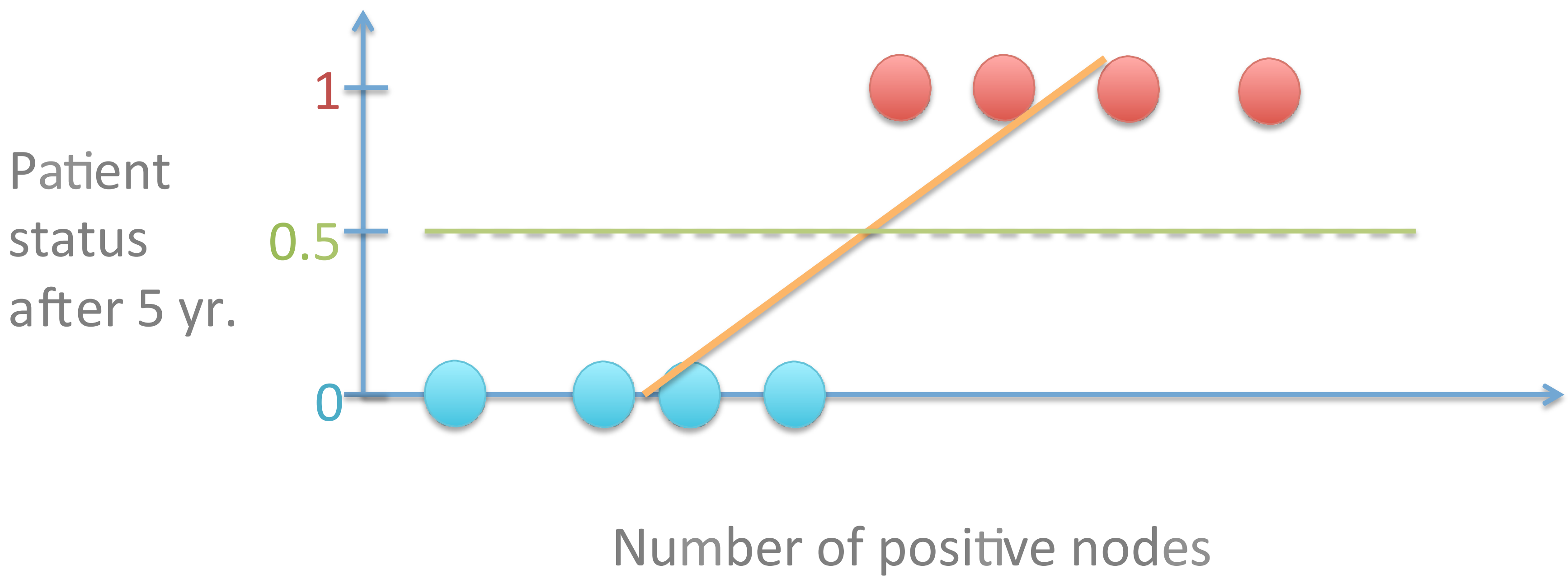


Linear regression for classification?



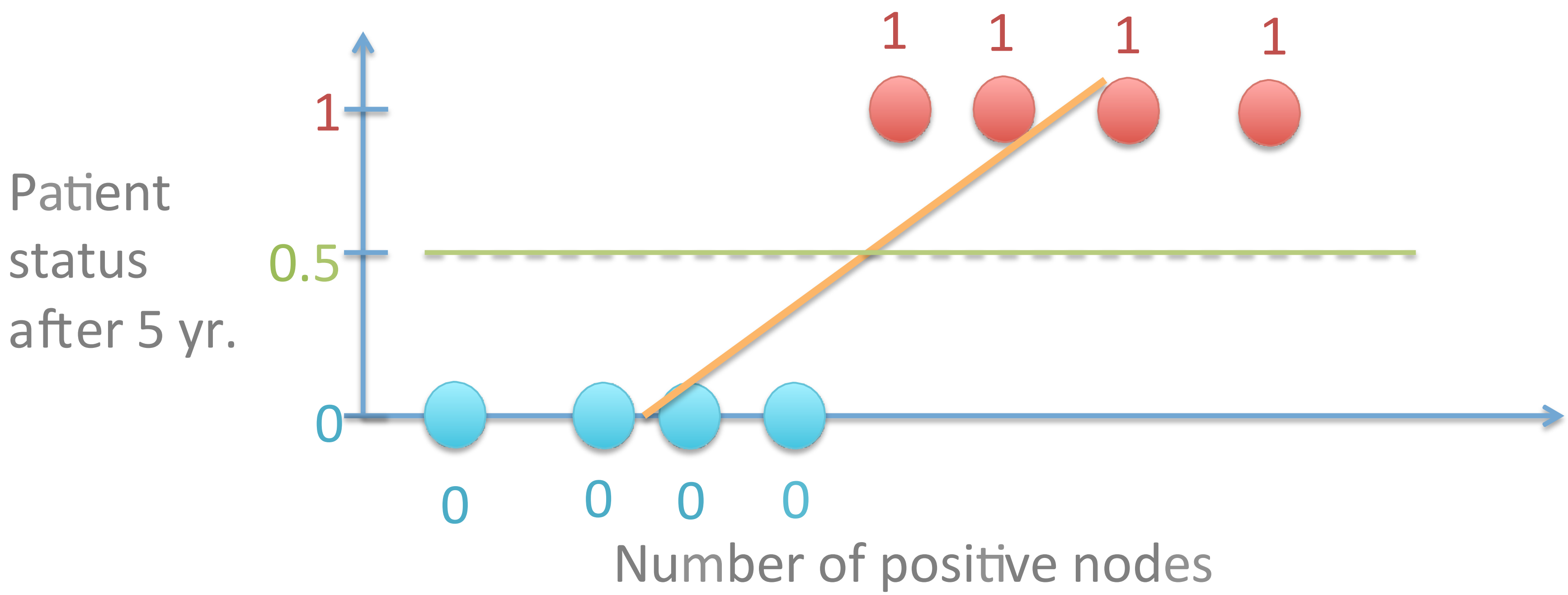
$$y_{\beta}(x) = \beta_0 + \beta_1 x + \varepsilon$$

Linear regression for classification?



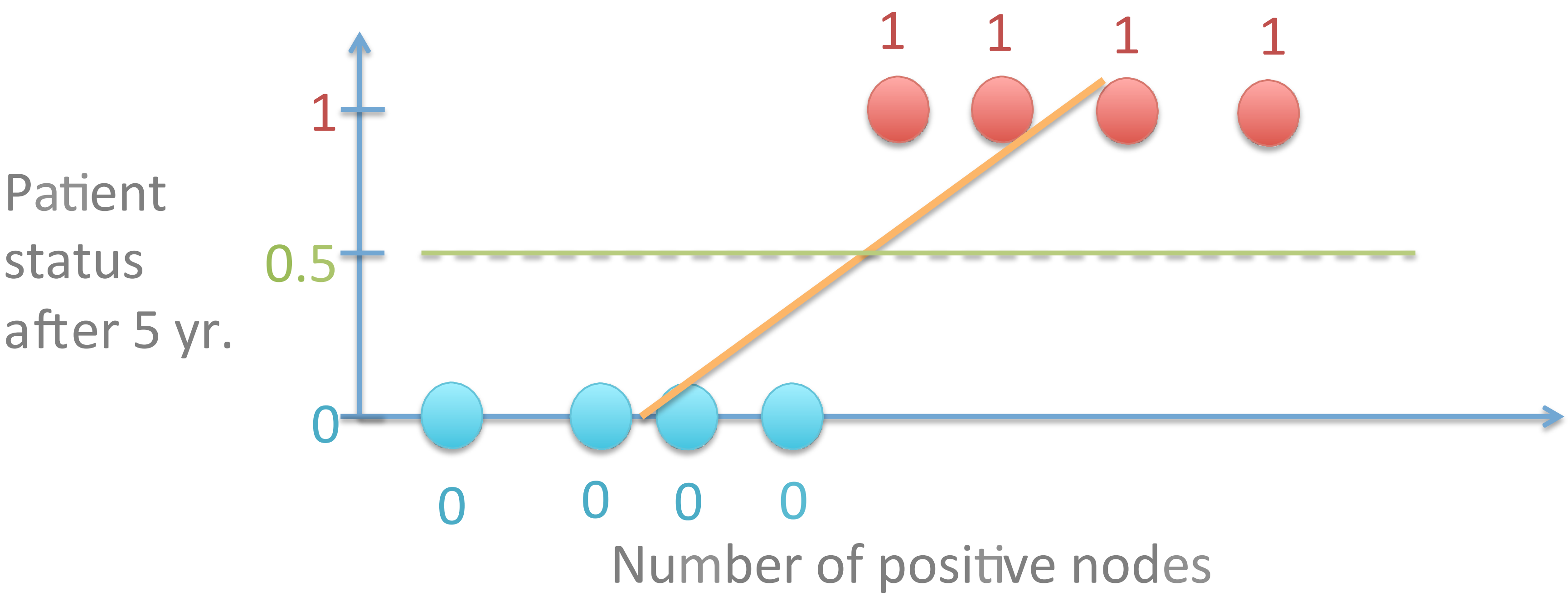
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



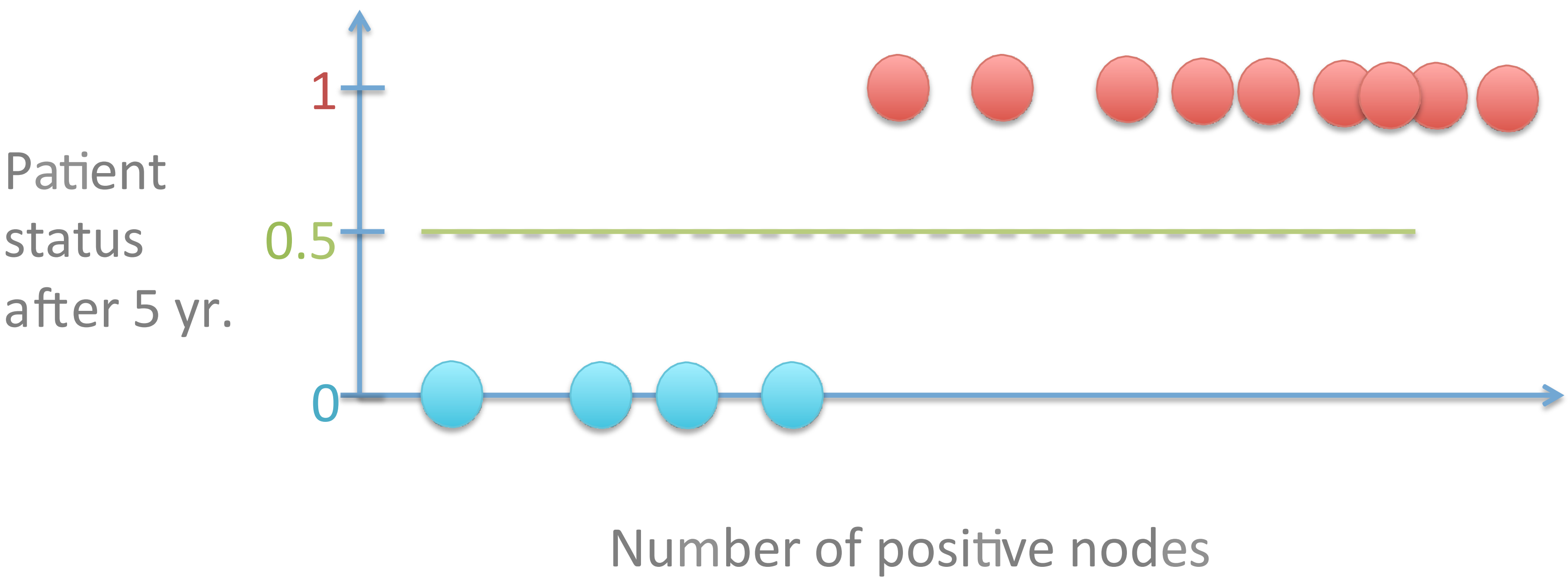
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



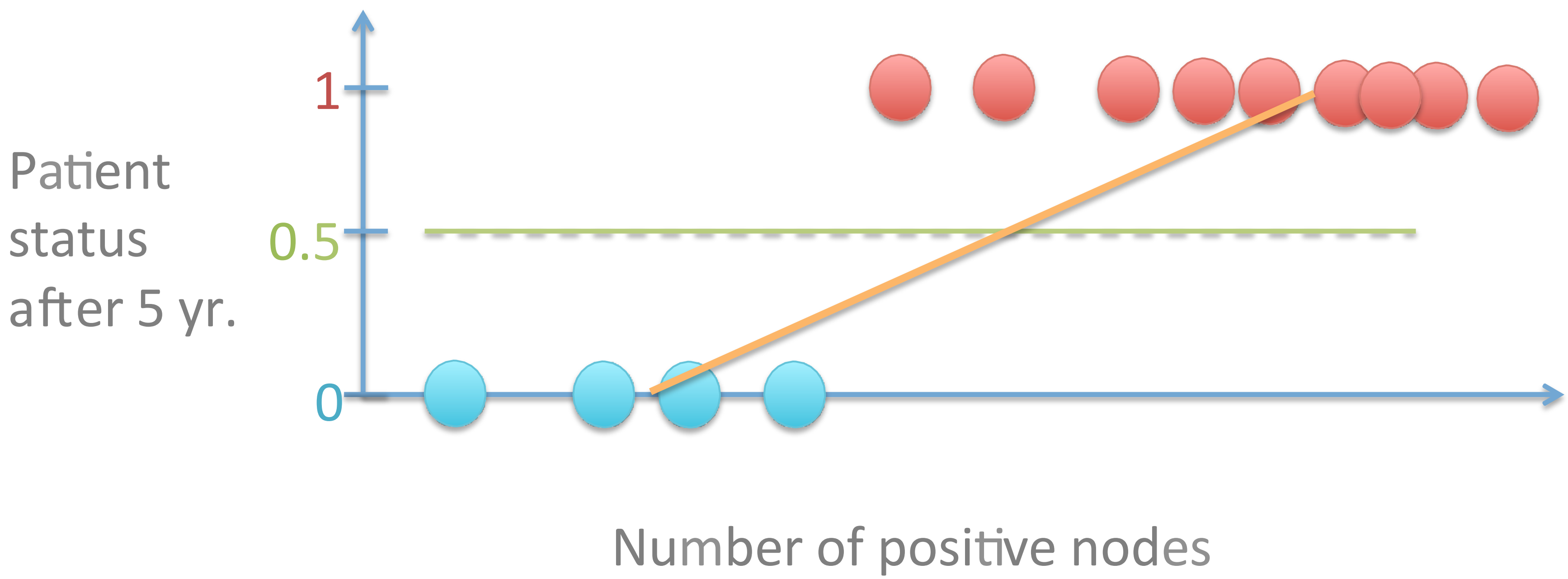
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



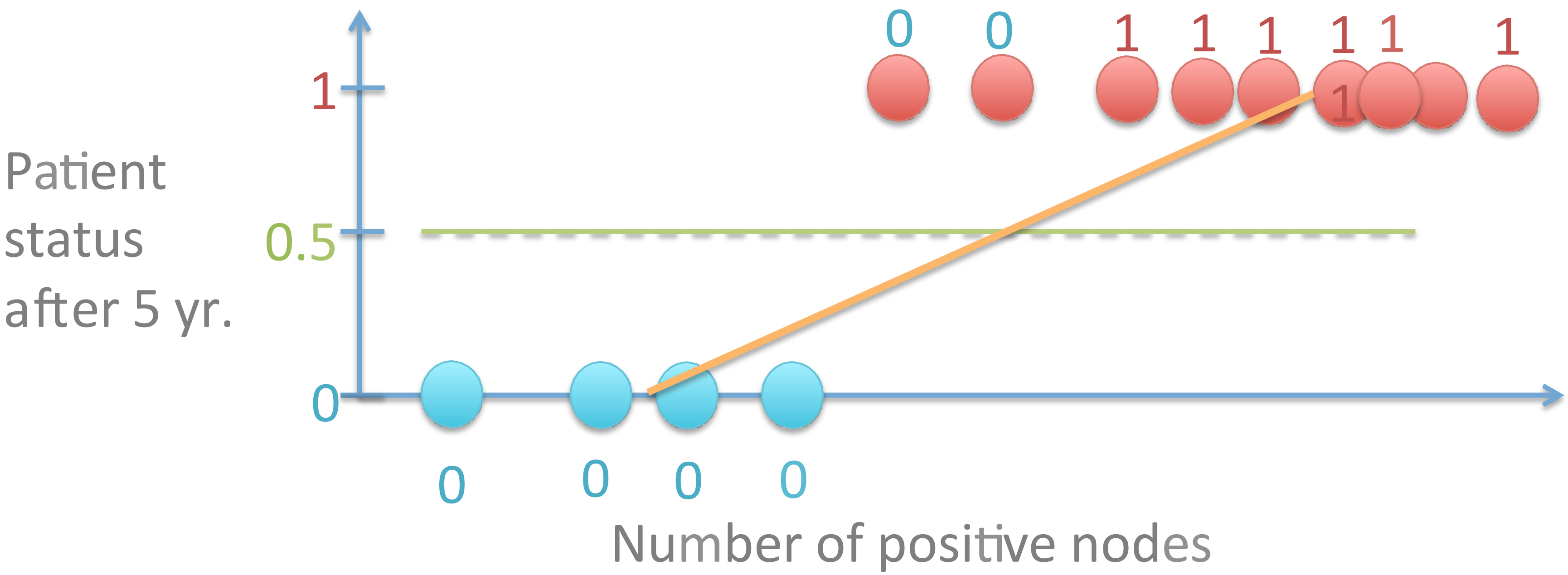
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



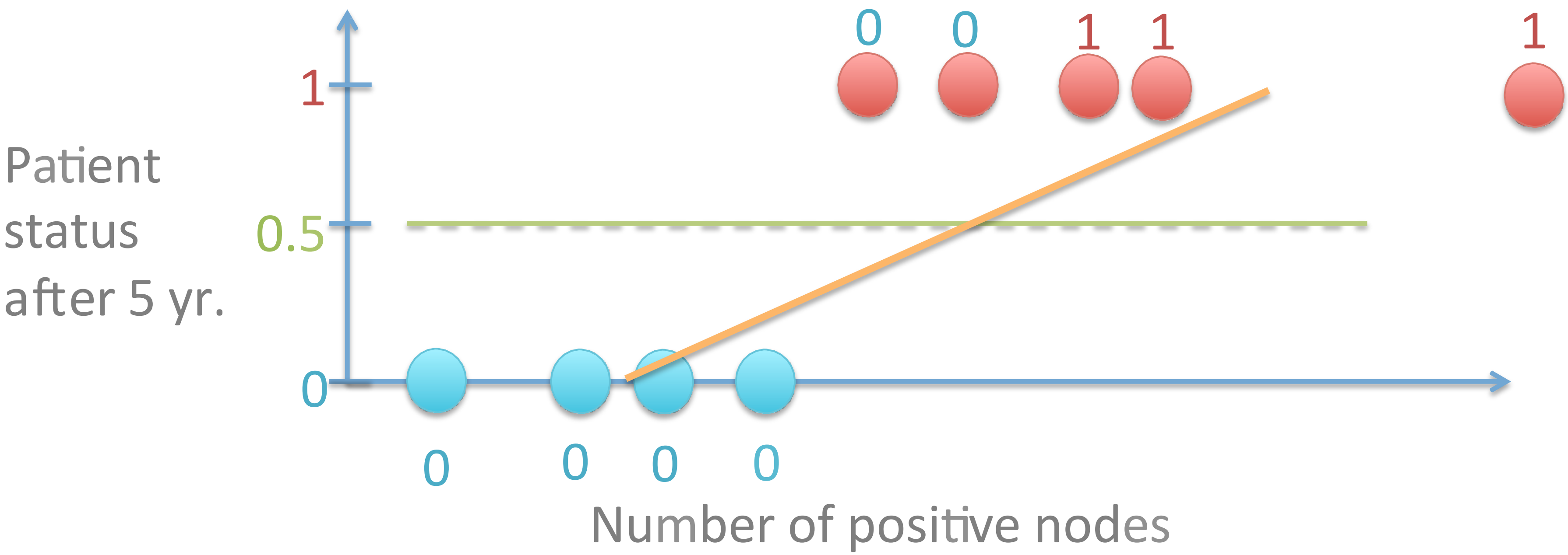
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



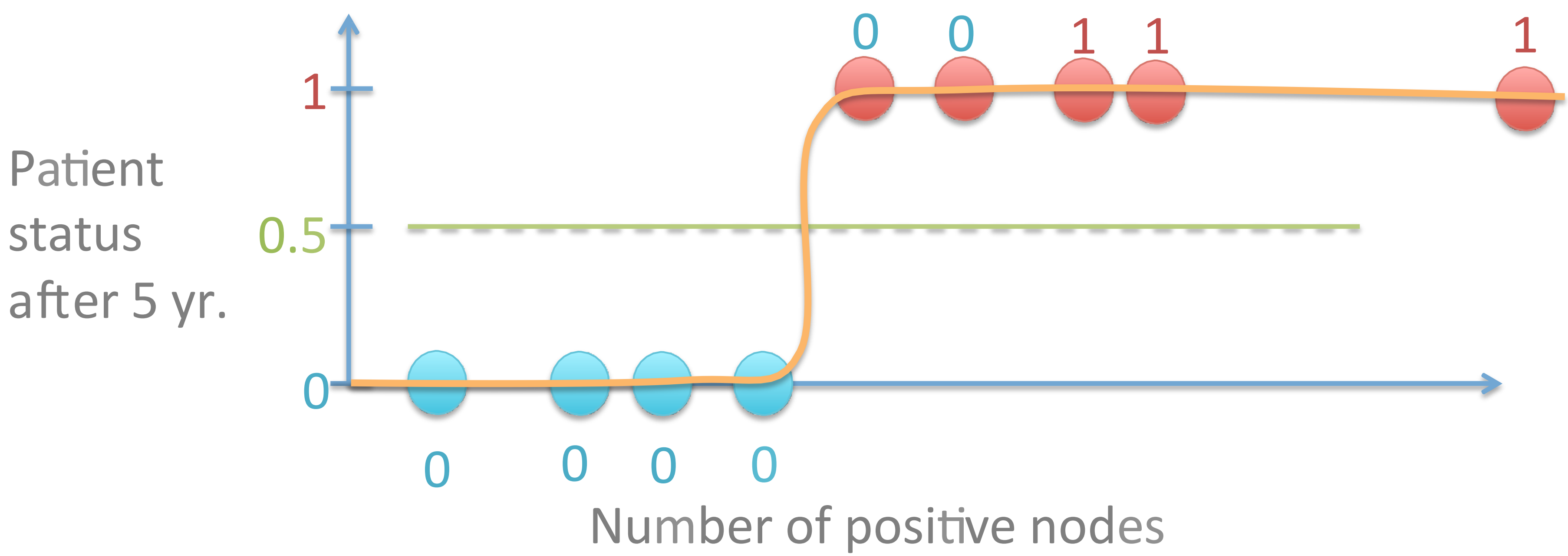
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Linear regression for classification?



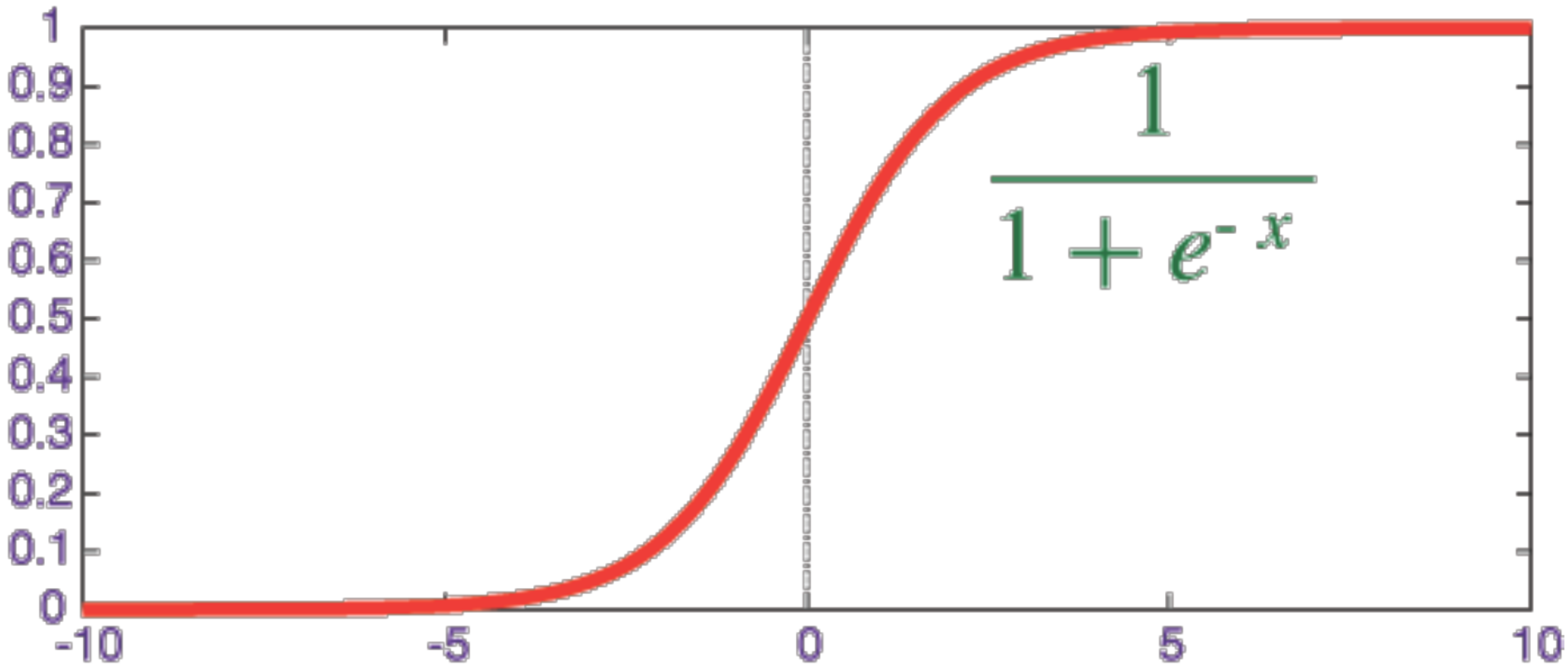
If y_pred > 0.5 predict label 1 (lost)
If y_pred < 0.5 predict label 0 (survived)

Logistic regression to the rescue

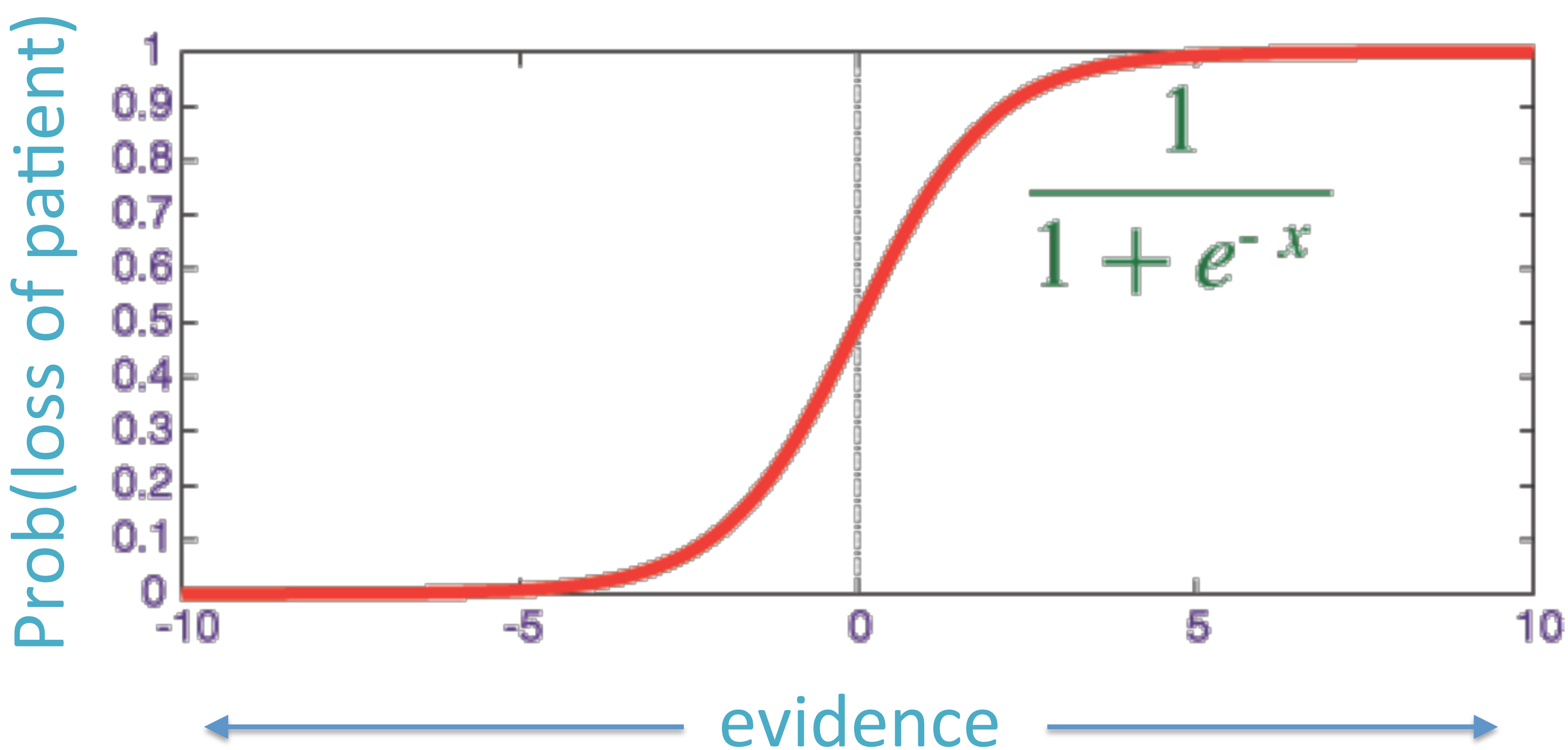


$$y_{\beta}(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x + \varepsilon)}}$$

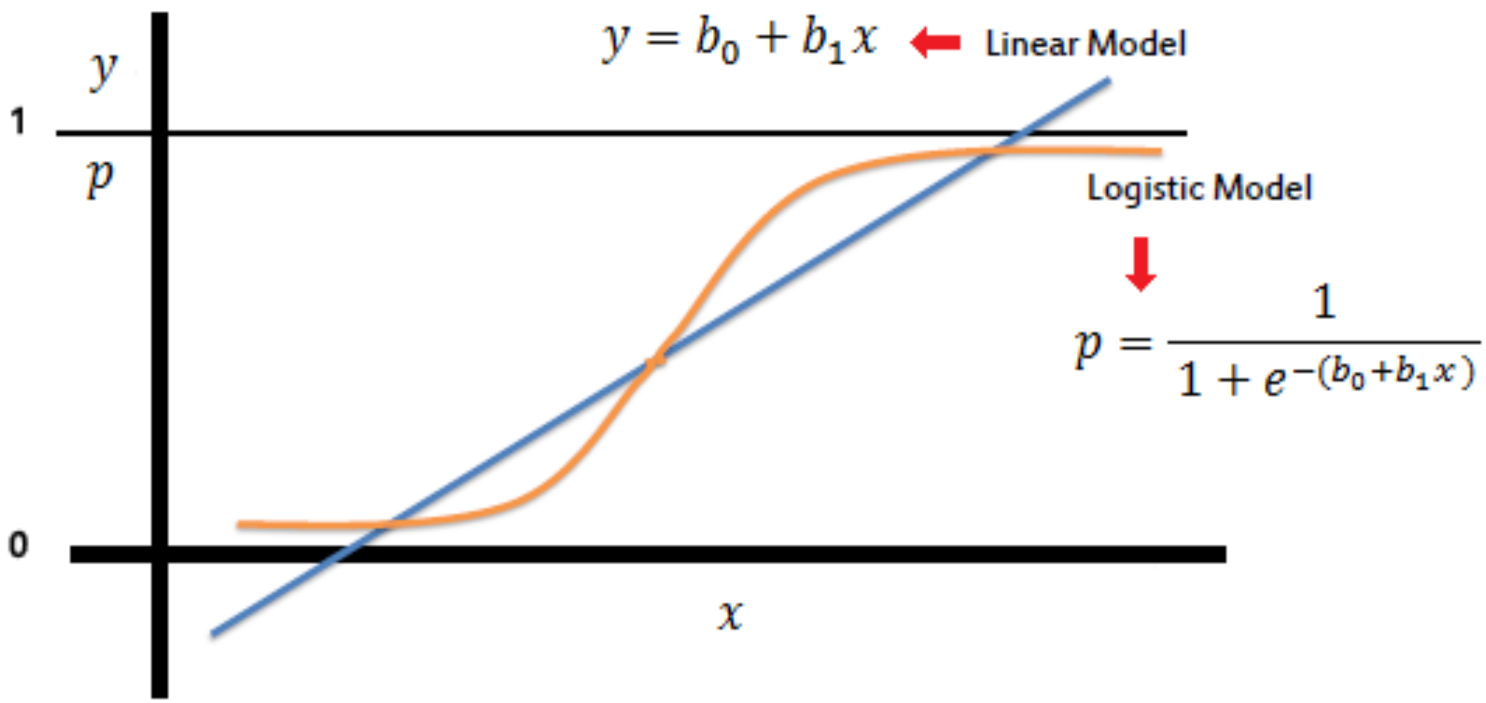
What is this function?



What is this function?

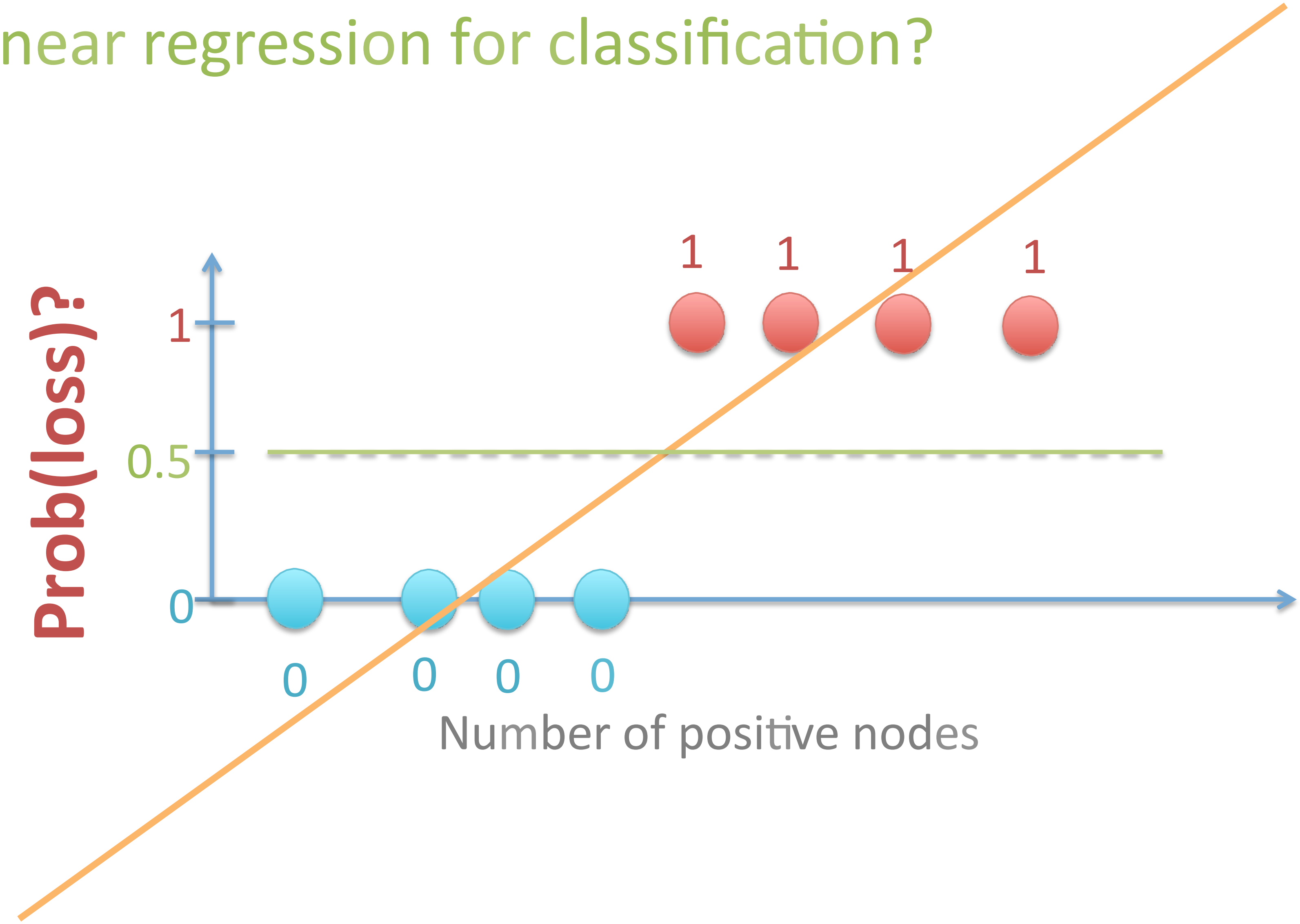


What is this function?

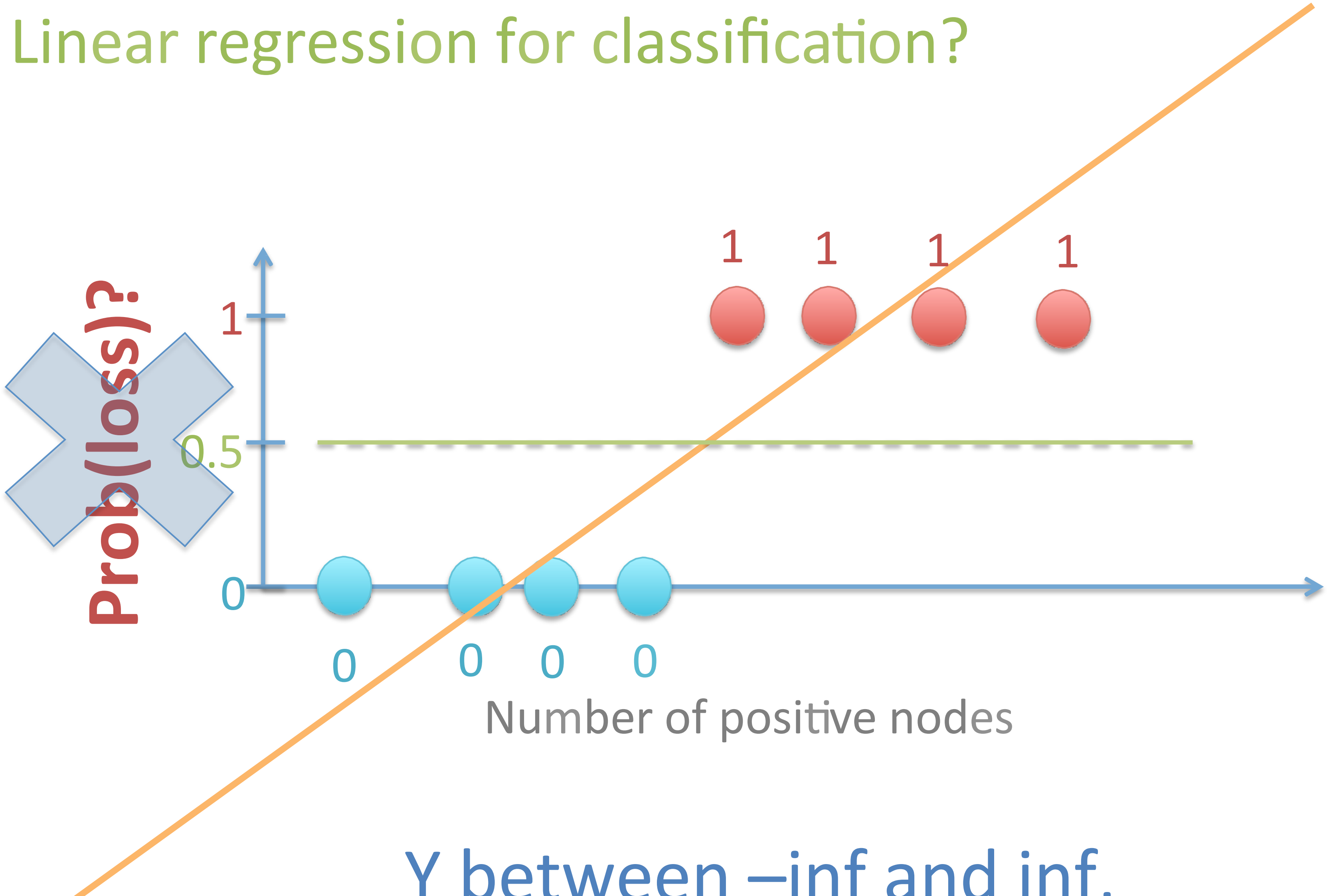


http://www.saedsayad.com/logistic_regression.htm

Linear regression for classification?



Linear regression for classification?



Y between –inf and inf,
Not 0 and 1

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.8$$

$$P(survival) = 0.2$$

Probability

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.8$$

$$P(survival) = 0.2$$

$$\frac{P(loss)}{P(survival)} = 4$$

Probability

Odds

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.05$$

$$P(survival) = 0.95$$

$$\frac{P(loss)}{P(survival)} = 0.053$$

Probability

Odds

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.5$$

$$P(survival) = 0.5$$

$$\frac{P(loss)}{P(survival)} = 1$$

Probability

Odds

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.5$$

$$P(survival) = 0.5$$

Probability

$$\frac{P(loss)}{P(survival)} = 1$$

Odds

between 0 and inf

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.5$$

$$P(survival) = 0.5$$

Probability

$$\log\left(\frac{P(loss)}{P(survival)}\right) = 0$$

Log Odds

between -inf and inf

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.05$$

$$P(survival) = 0.95$$

$$\log\left(\frac{P(loss)}{P(survival)}\right) = -2.94$$

Probability

Log Odds

between -inf and inf

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.8$$

$$P(survival) = 0.2$$

$$\log\left(\frac{P(loss)}{P(survival)}\right) = 1.39$$

Probability

Log Odds

between -inf and inf

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.999$$
$$P(survival) = 0.001$$

$$\log\left(\frac{P(loss)}{P(survival)}\right) = 6.9$$

Probability

Log Odds

between -inf and inf

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.999$$

$$1 - P(loss) = 0.001$$

$$\log\left(\frac{P(loss)}{1 - P(loss)}\right) = 6.9$$

Probability

Log Odds

logit function

What metric would express the chances of loss/survival, but not constrained to [0, 1] ?

$$P(loss) = 0.999$$

$$1 - P(loss) = 0.001$$

Probability

$$\log\left(\frac{P(loss)}{1 - P(loss)}\right) = 6.9$$

Log Odds

logit function

$$\frac{1}{1 + e^{-\log\left(\frac{P(loss)}{1-P(loss)}\right)}} = P(loss)$$

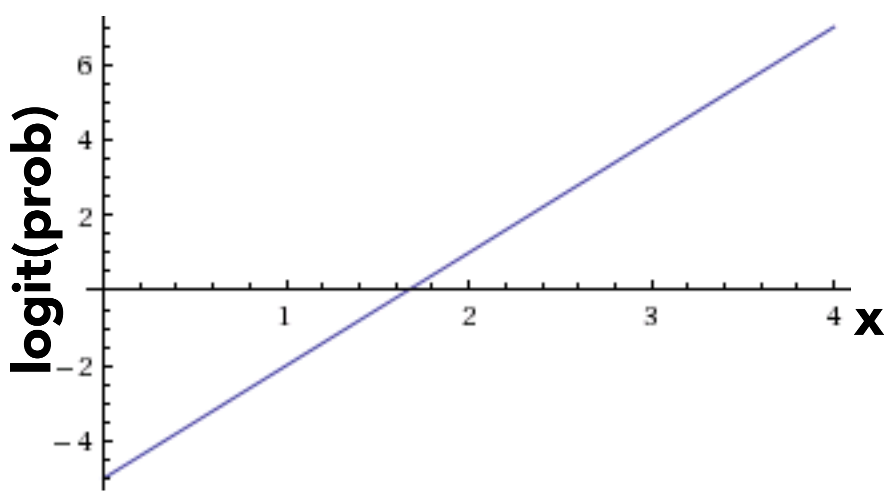
Logistic Function

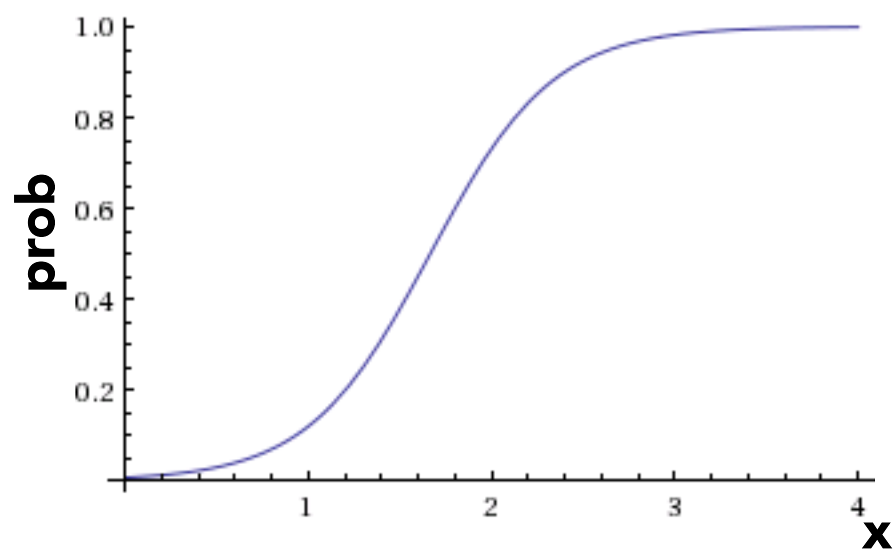
Log Odds → Prob

What is this function?

logit(prob **)** = log odds

logistic(log odds **)** = prob





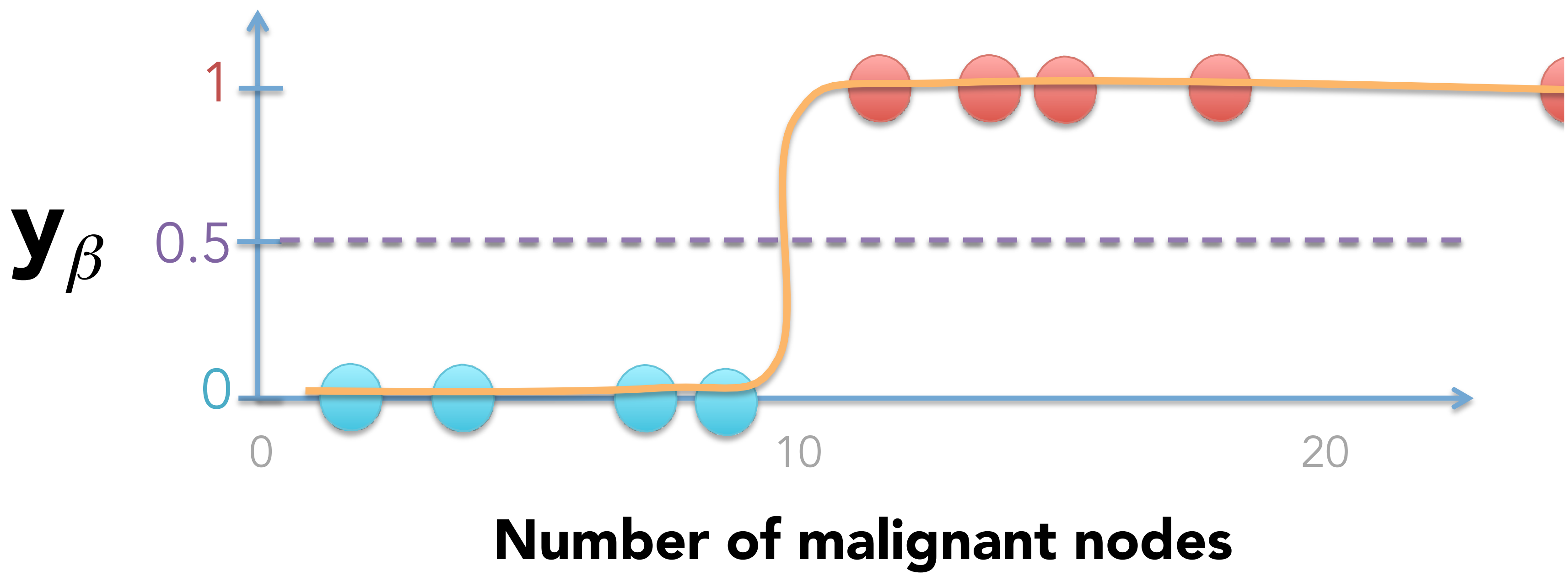
Coefficients work the same way

logit(prob**)** = log odds = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \varepsilon$

increase in log odds
per 1 unit of x_1

increase in log odds
per 1 unit of x_2

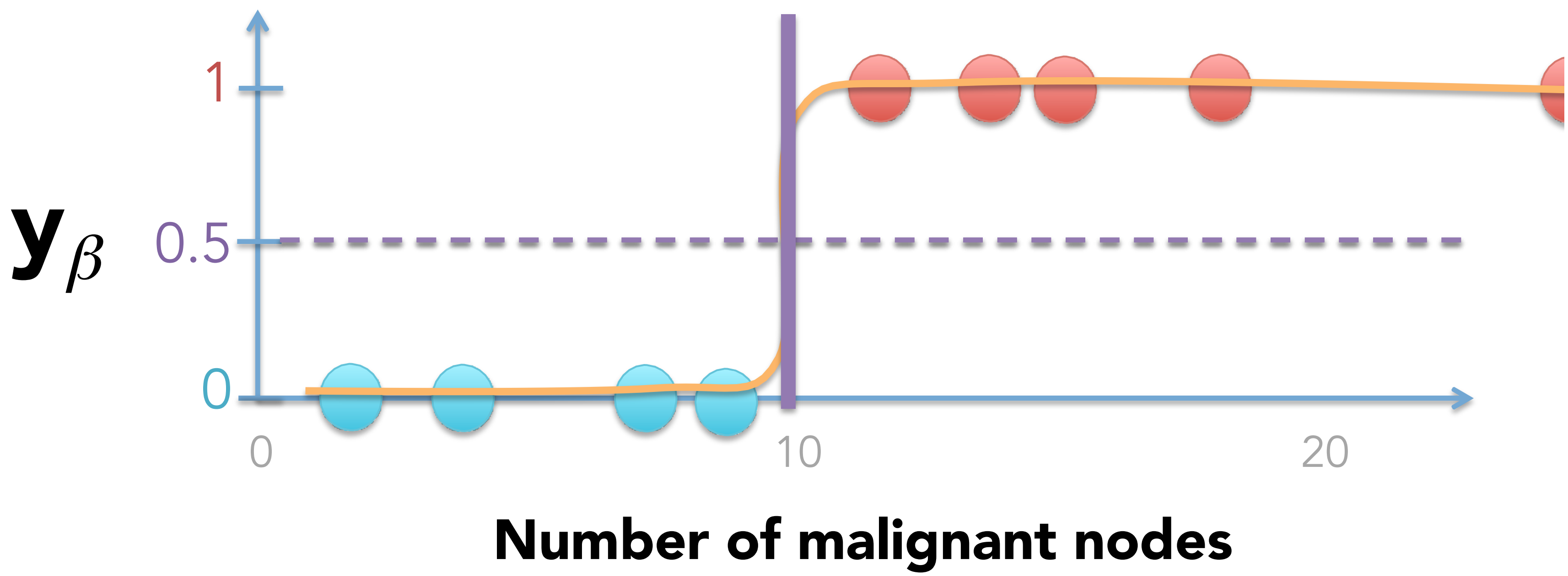
The "Decision Boundary"



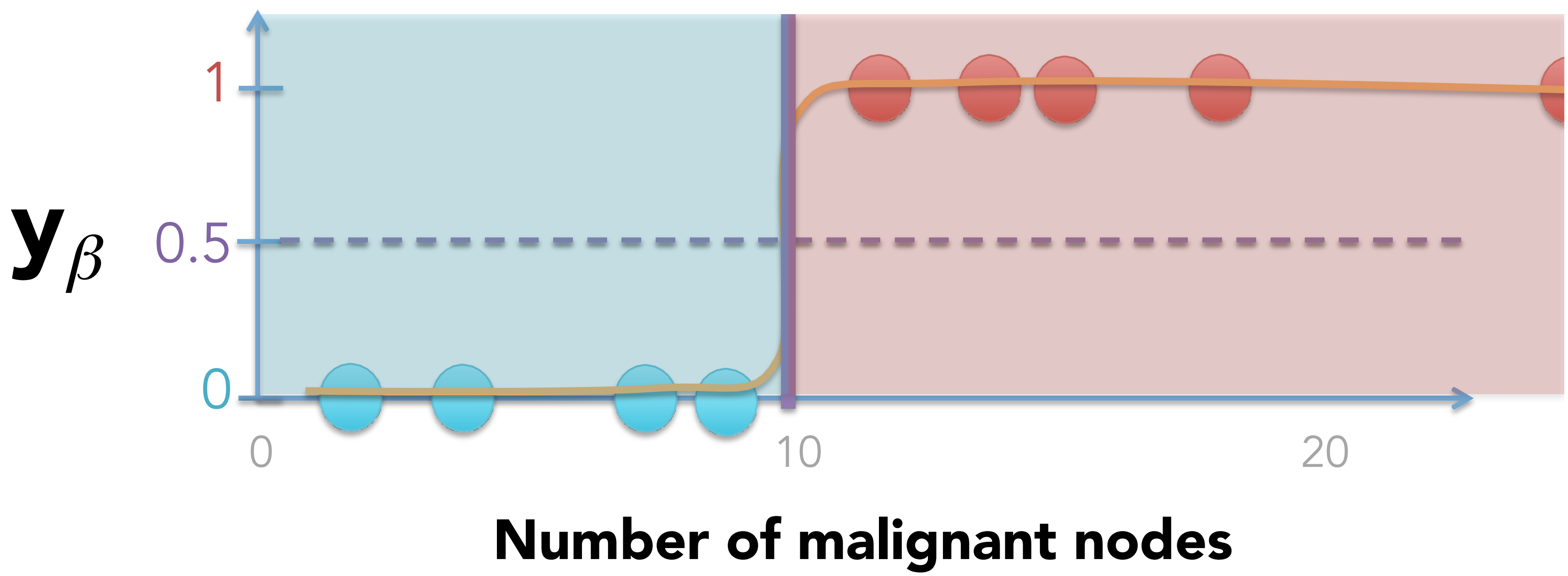
Predict 1 (lost) if $y_\beta > 0.5$

Predict 0 (survived) if $y_\beta < 0.5$

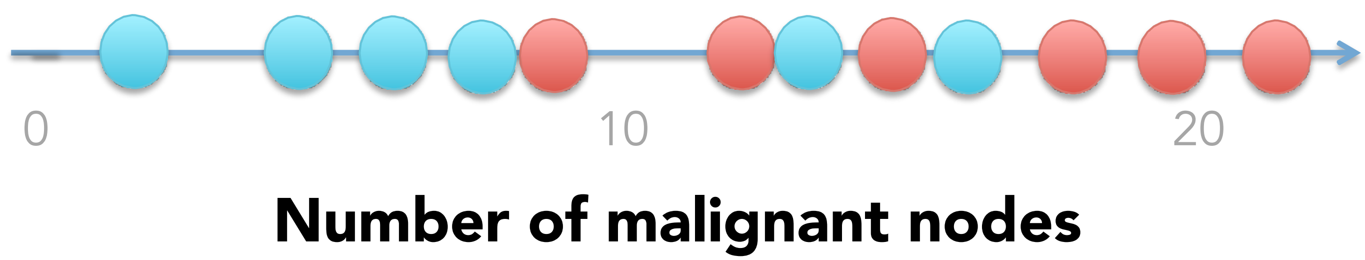
The "Decision Boundary"



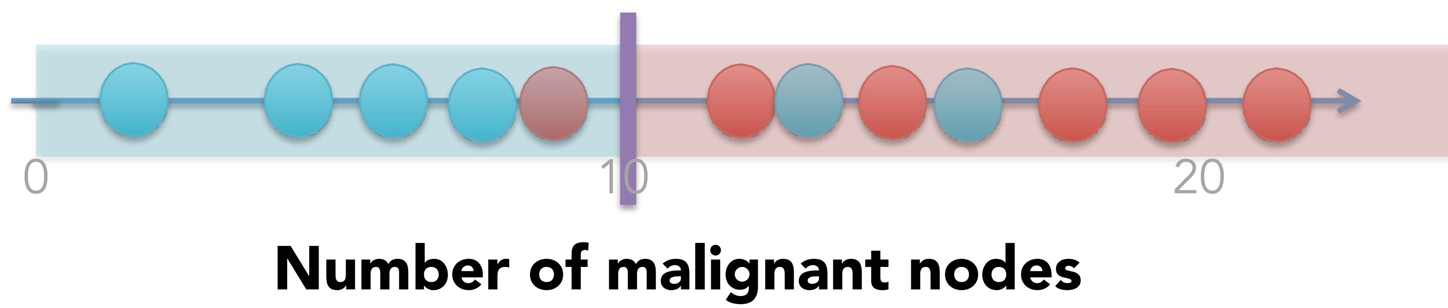
The "Decision Boundary"



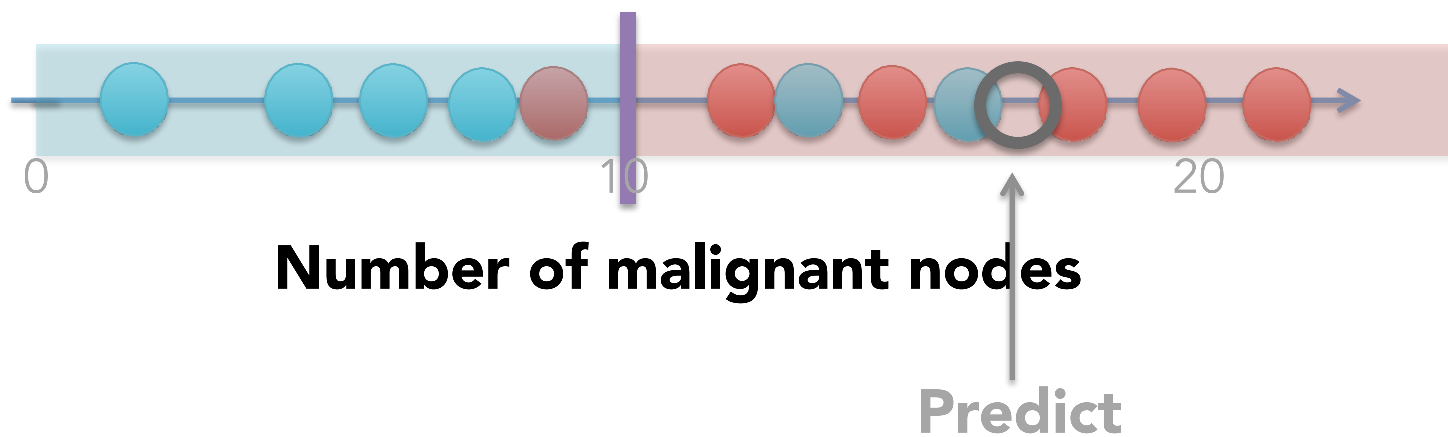
1 Feature. Number of malignant nodes
2 Labels. Survived / Lost



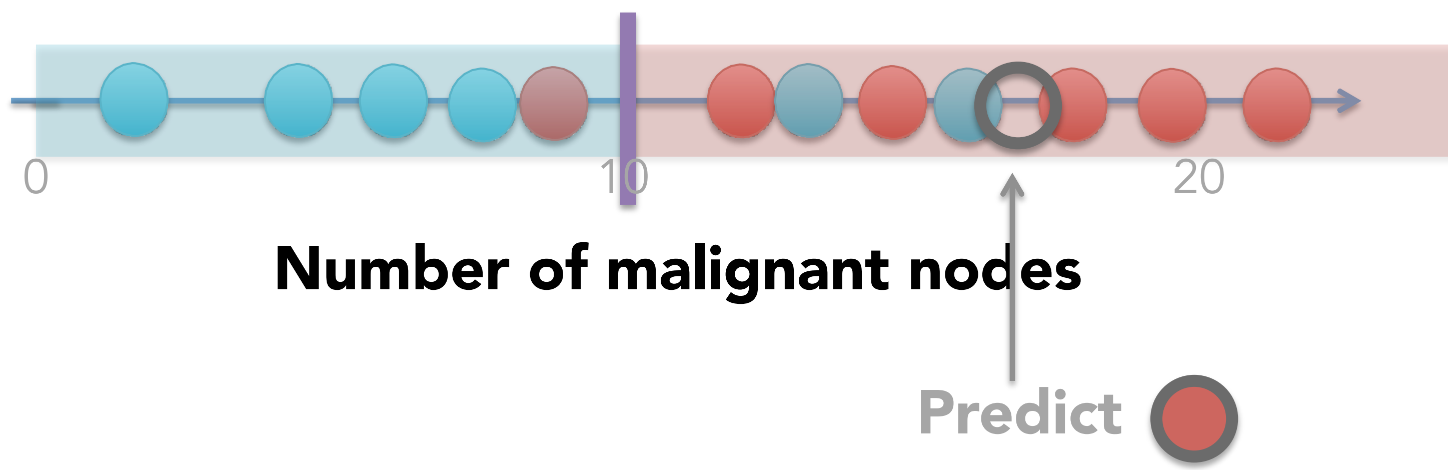
1 Feature. Number of malignant nodes
2 Labels. Survived / Lost



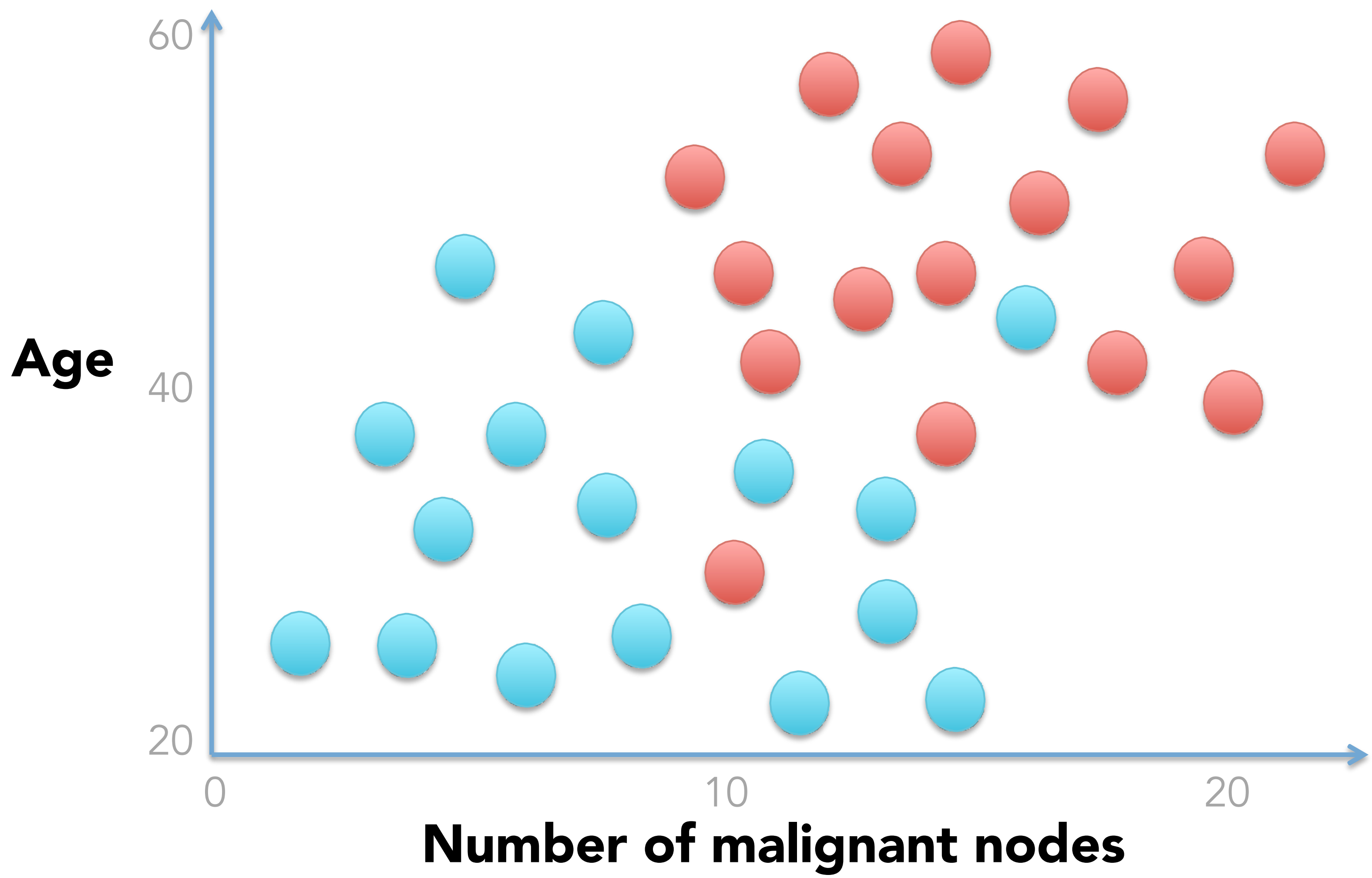
1 Feature. Number of malignant nodes
2 Labels. Survived / Lost



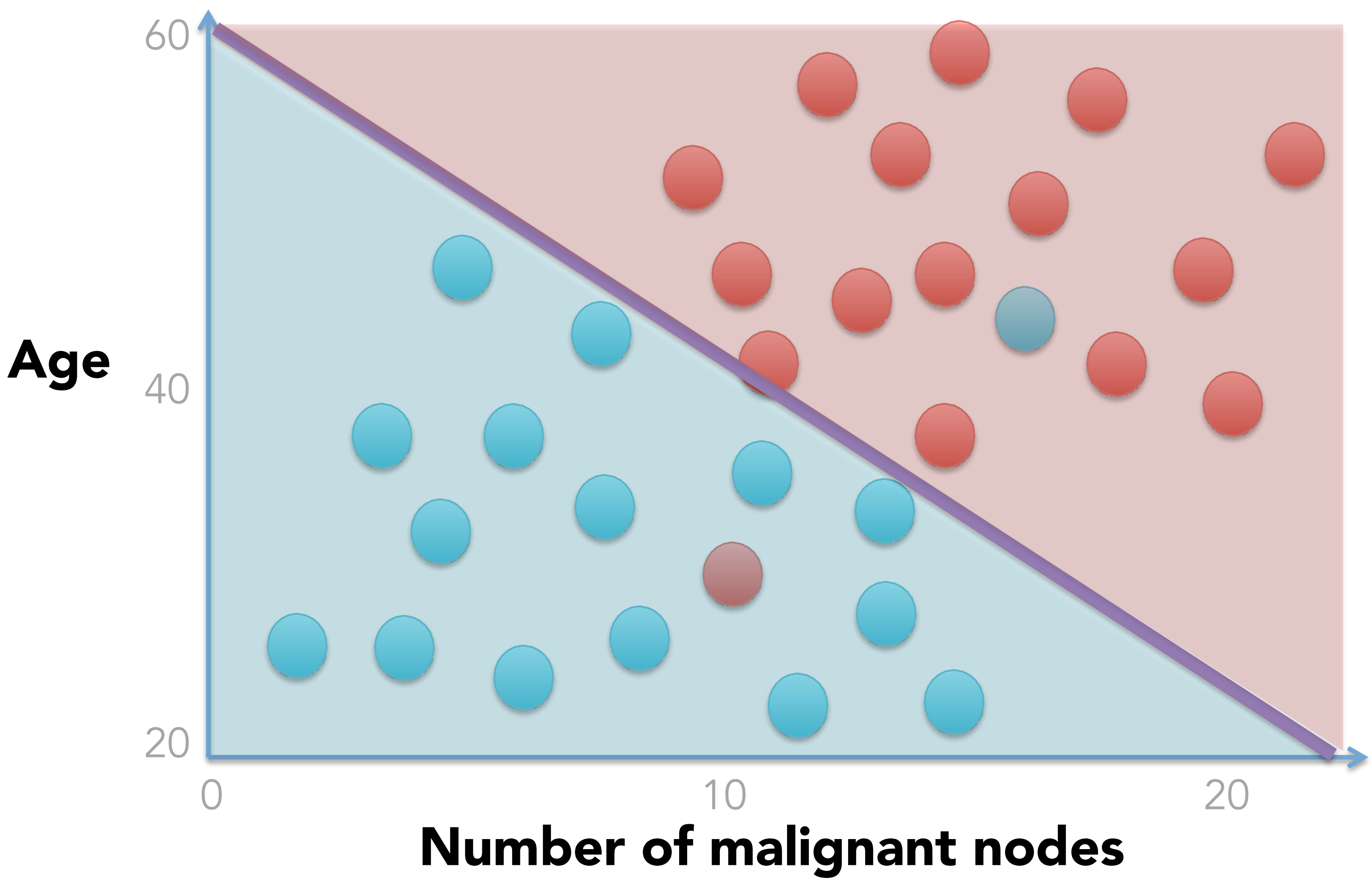
1 Feature. Number of malignant nodes
2 Labels. Survived / Lost



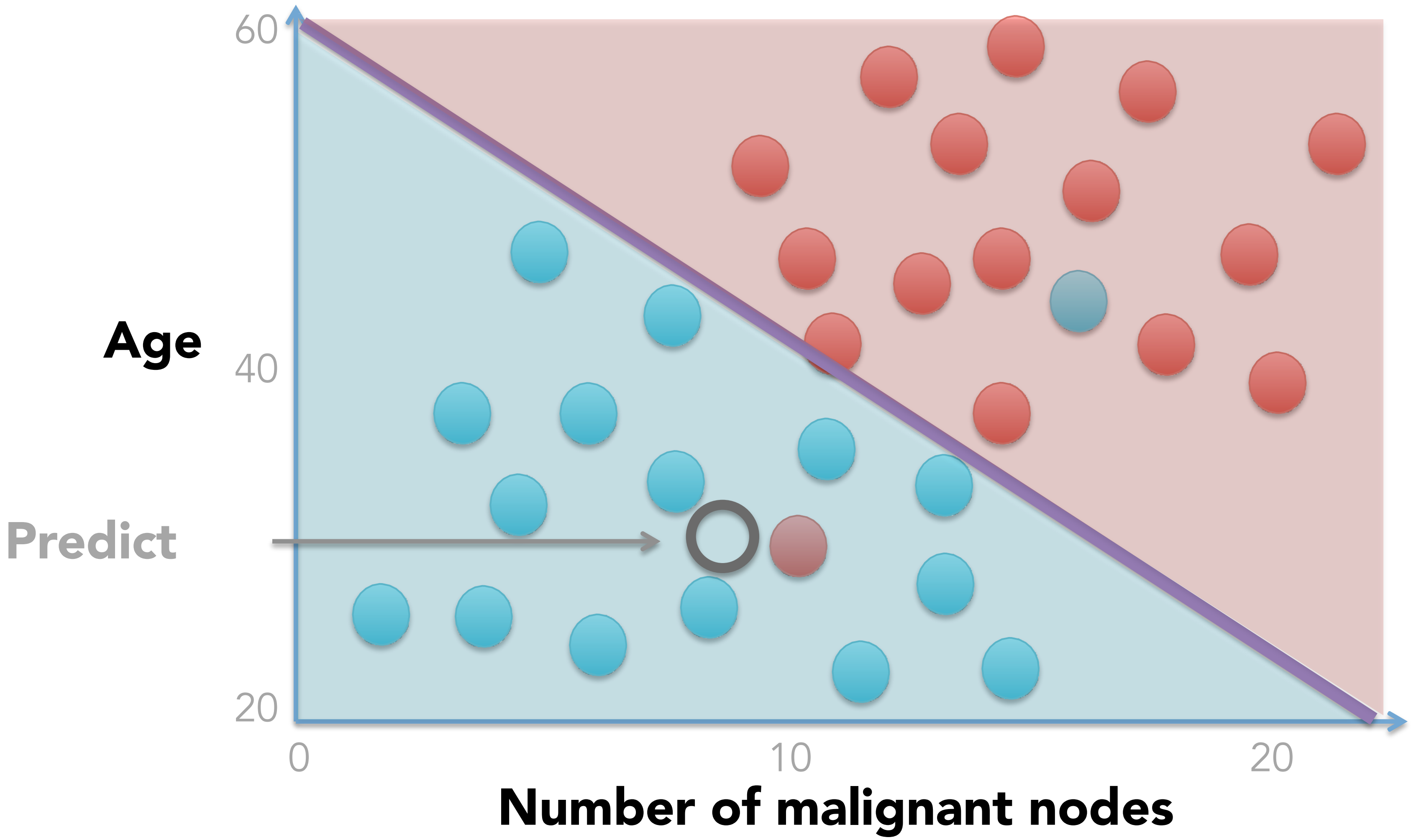
2 Features. No of malignant nodes / Age
2 Labels. Survived / Lost



2 Features. No of malignant nodes / Age
2 Labels. Survived / Lost

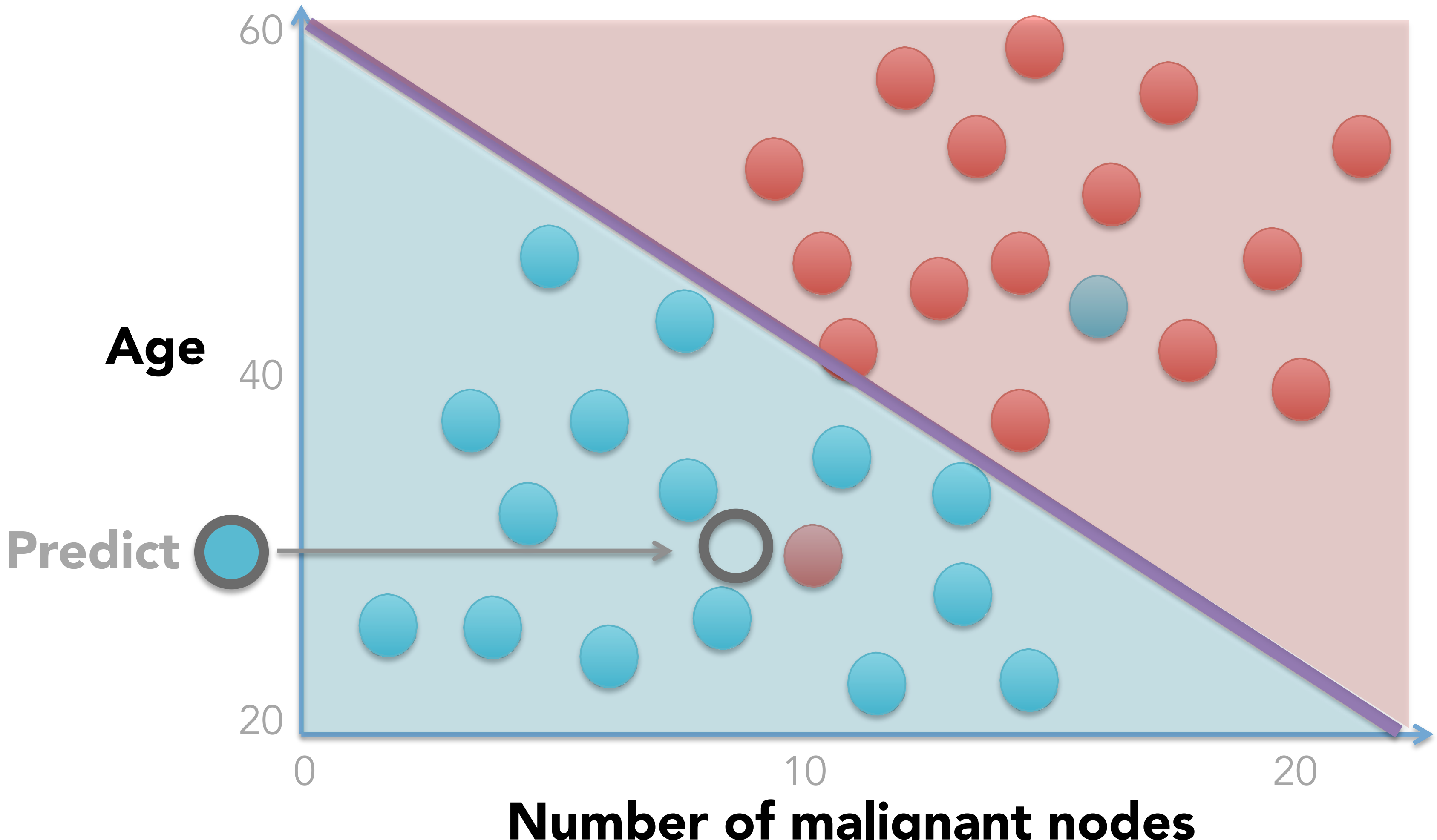


2 Features. No of malignant nodes / Age
2 Labels. Survived / Lost

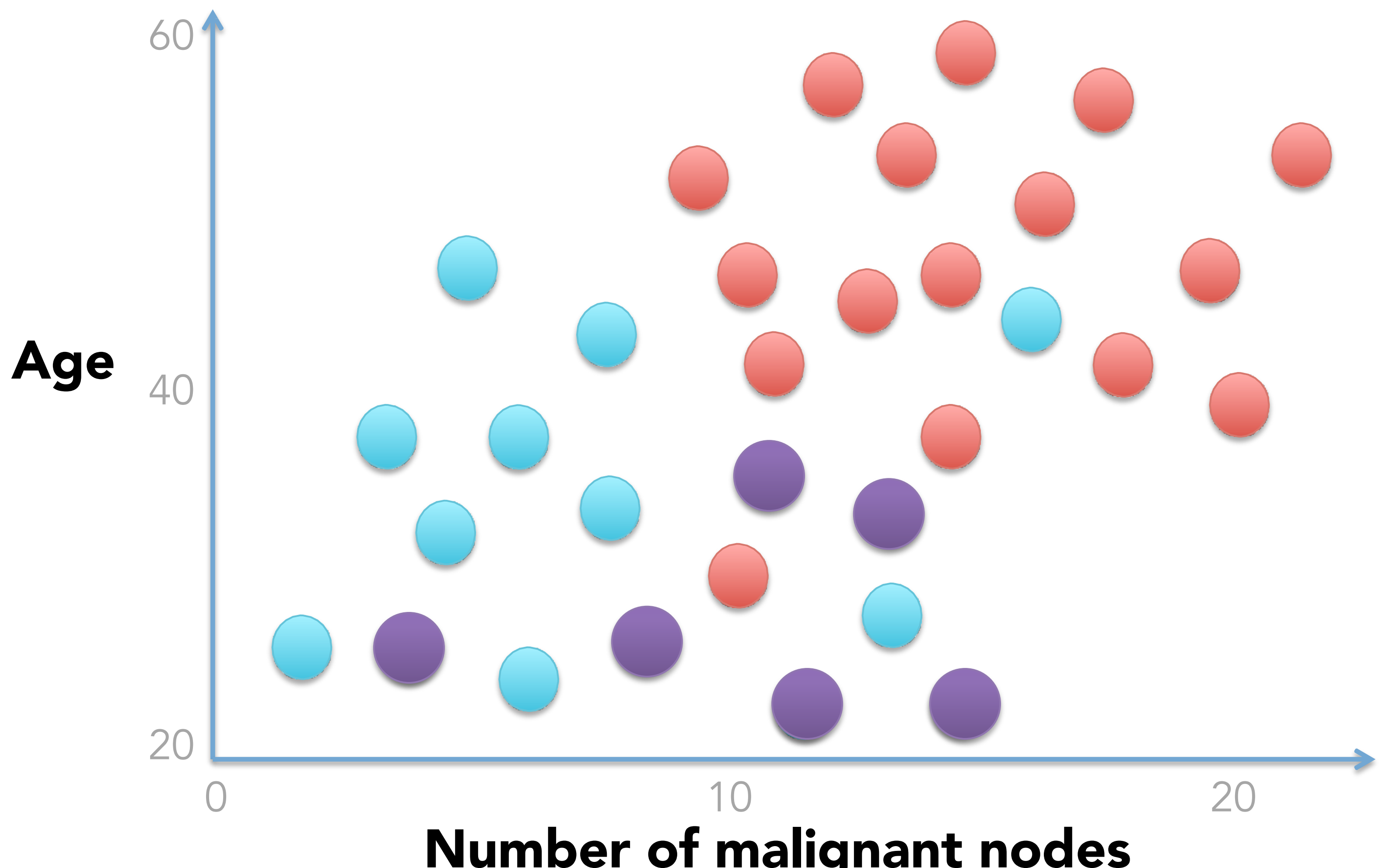


2 Features. No of malignant nodes / Age

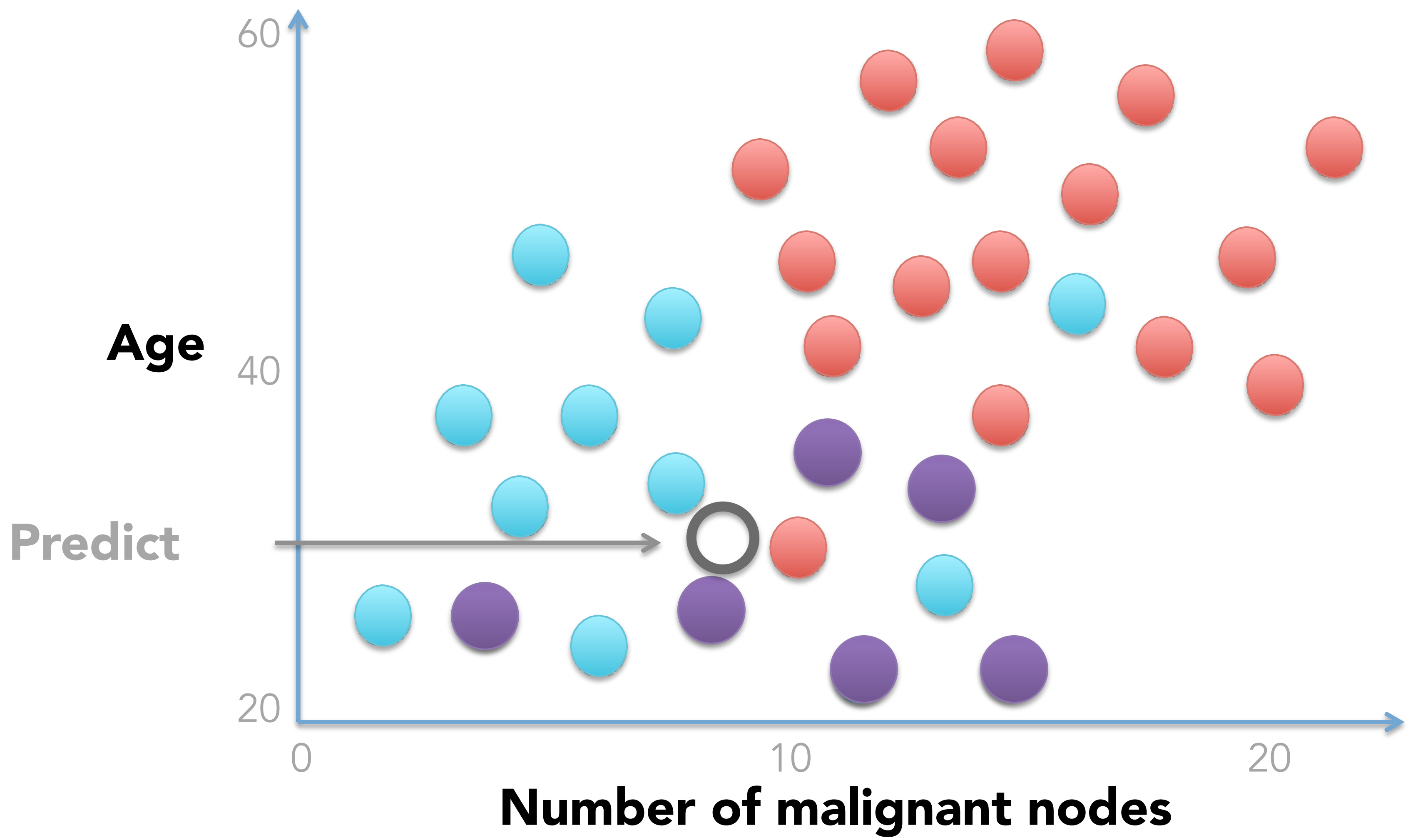
2 Labels. Survived / Lost



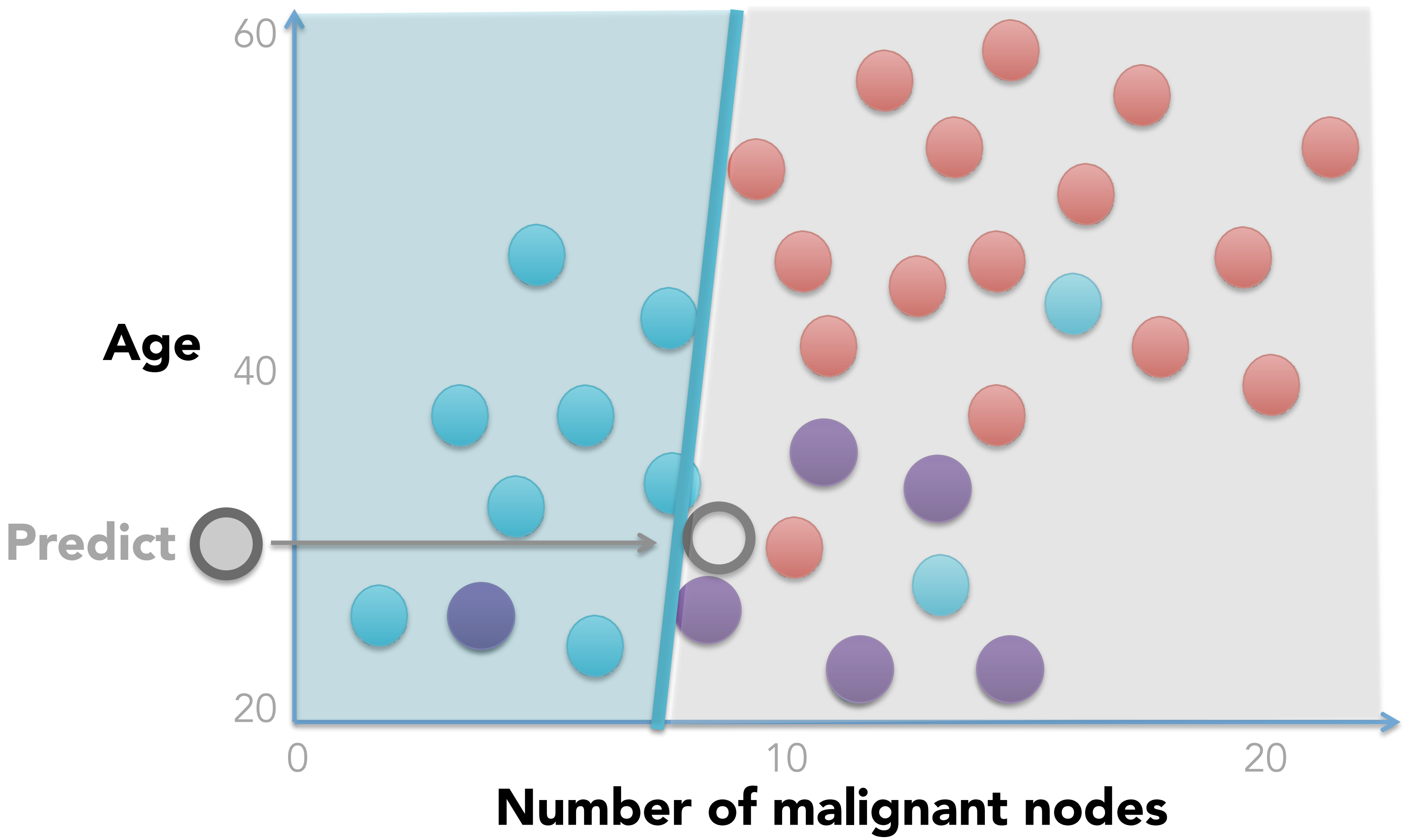
2 Features. No of malignant nodes / Age
3 Labels. Healthy / Complications / Lost



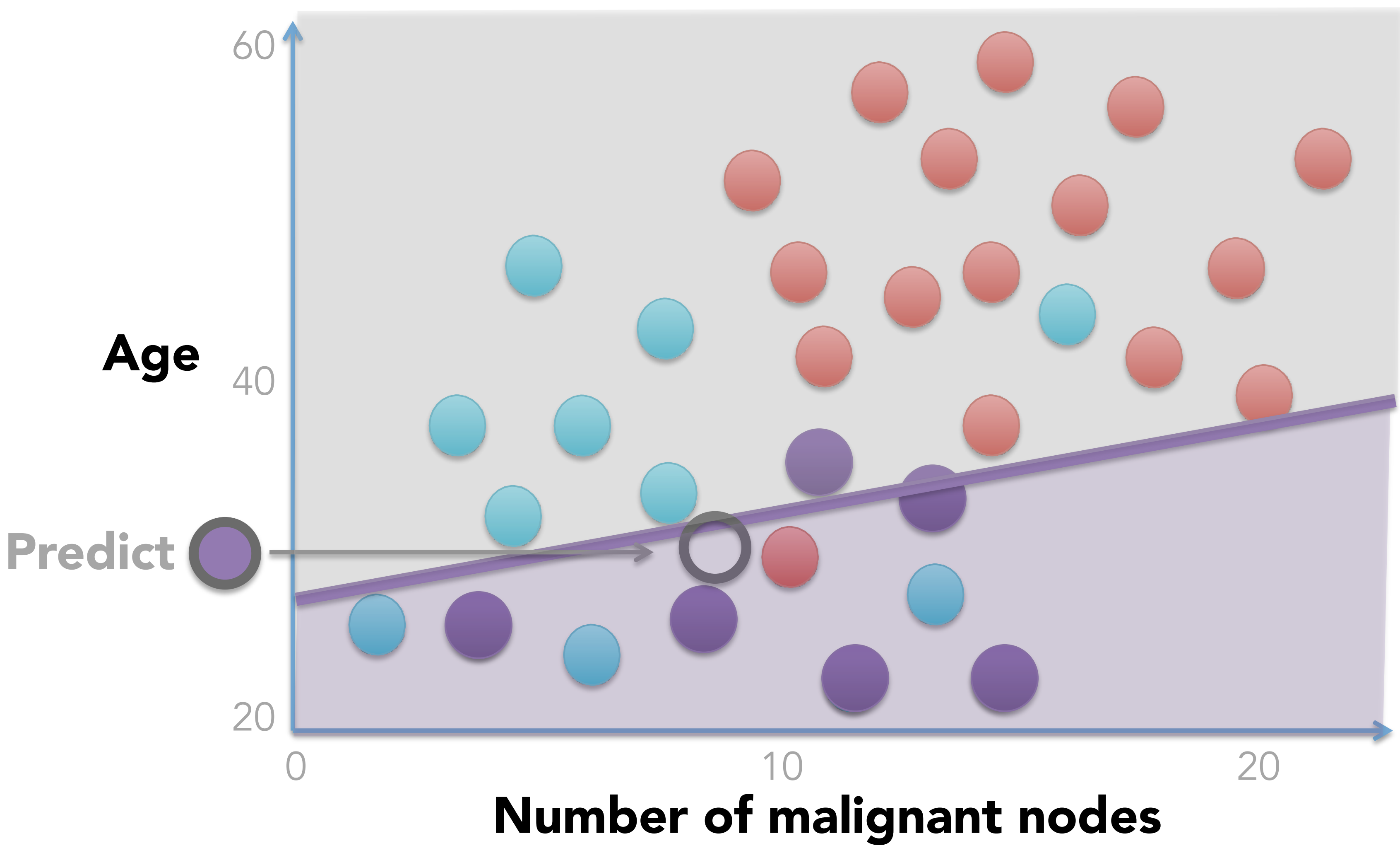
2 Features. No of malignant nodes / Age
3 Labels. Healthy / Complications / Lost



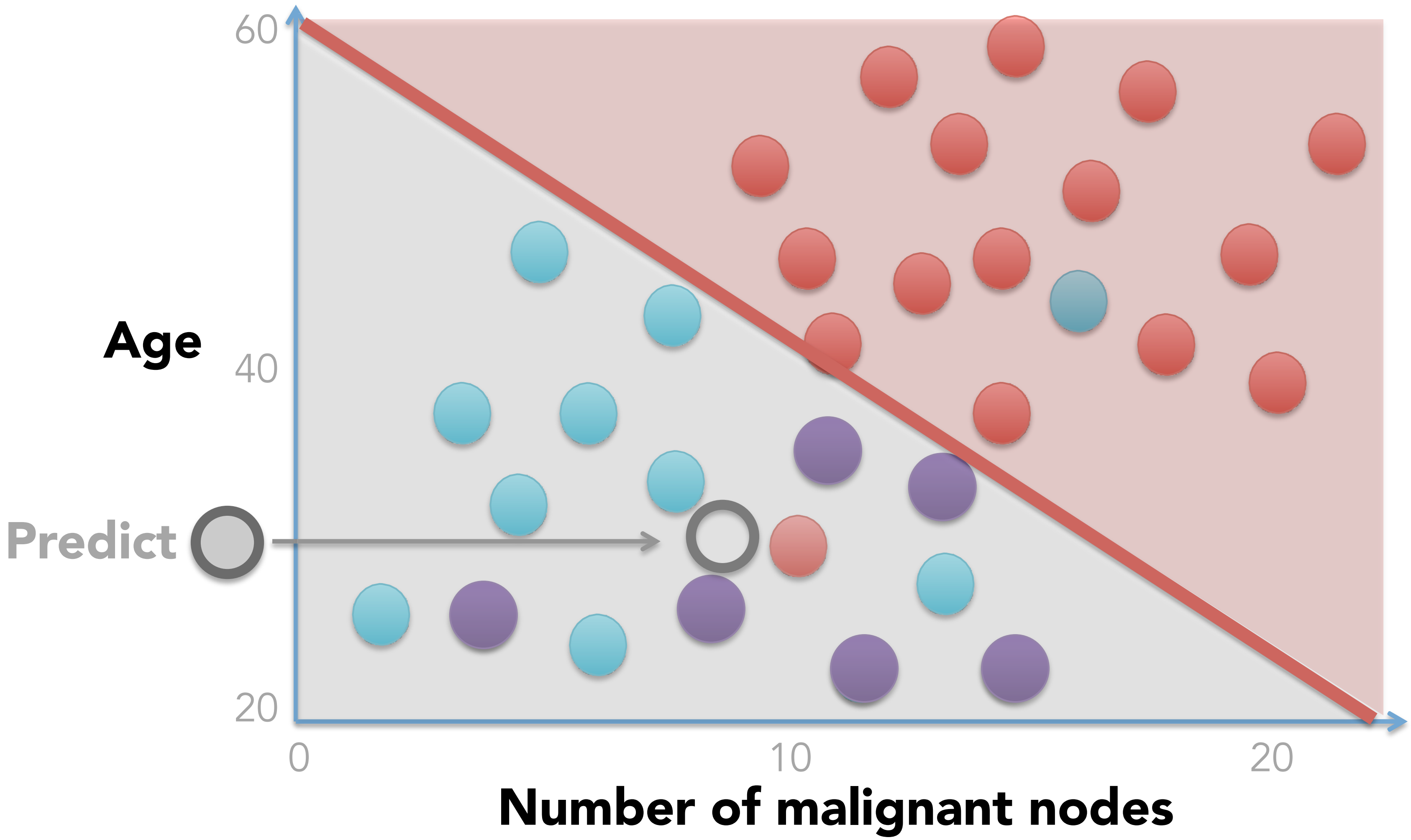
One vs all. Survived vs all.



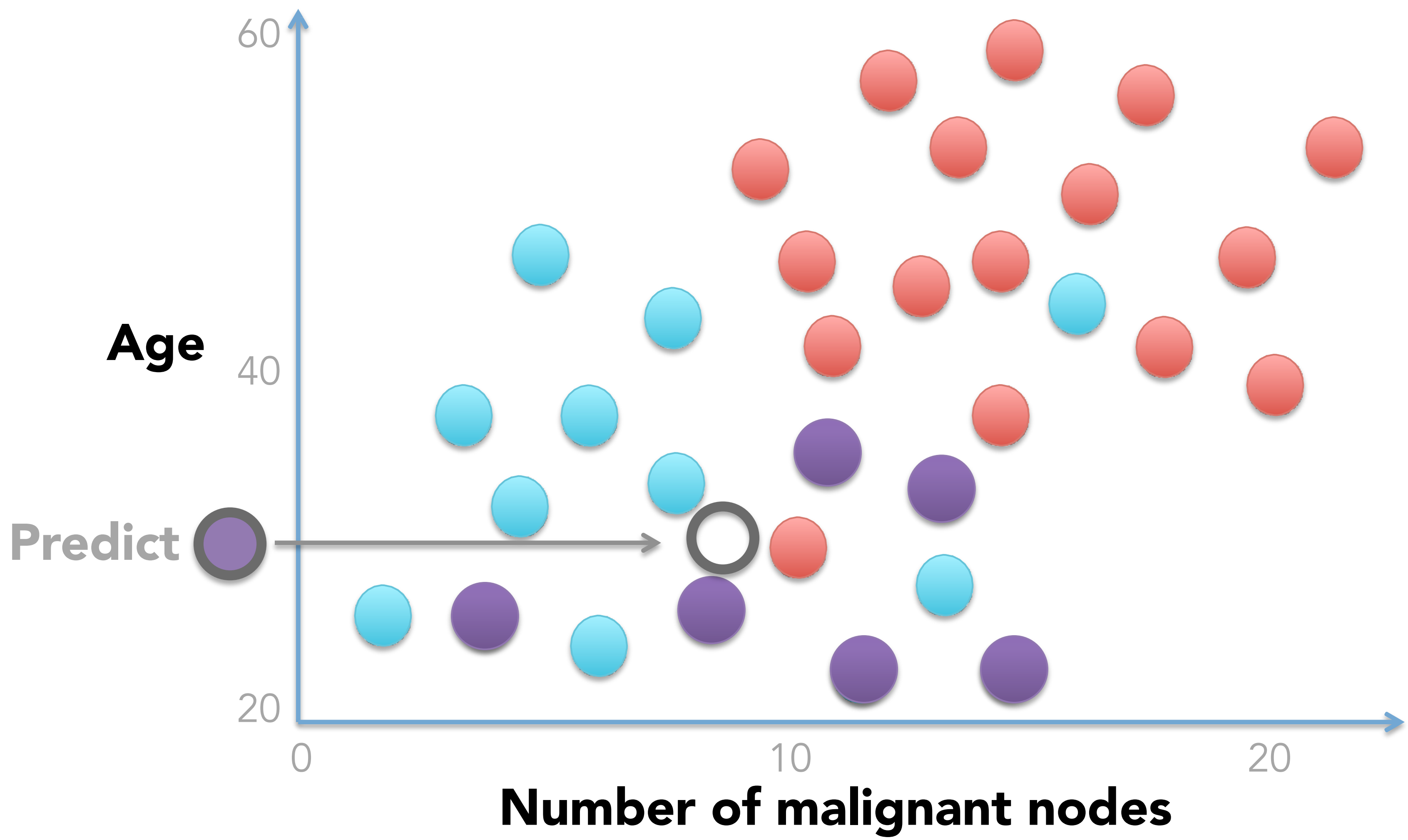
One vs all. Complications vs all.



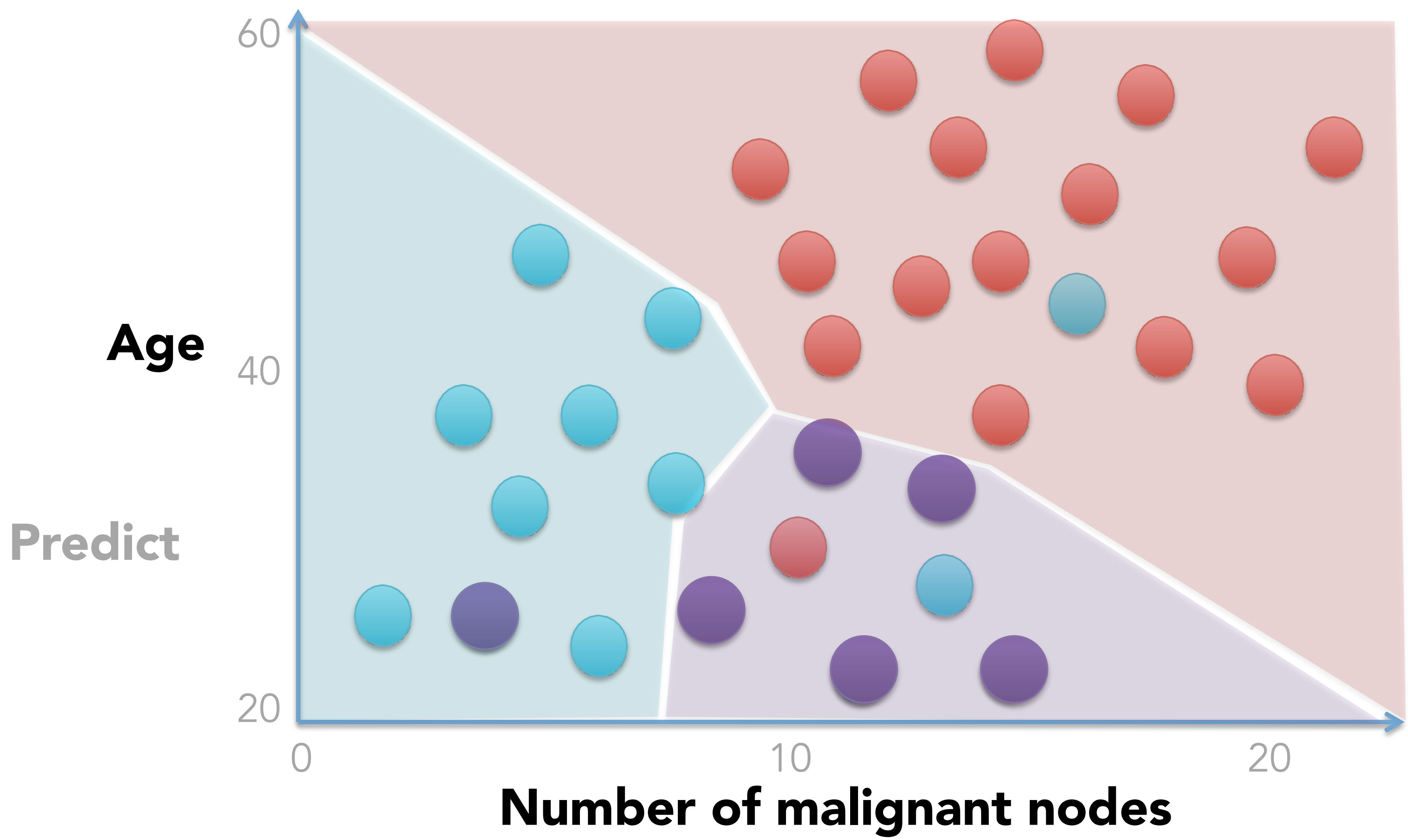
One vs all. Lost vs all.



One vs all. Winner: Complications



One vs all. Essentially, it becomes:



"Logistic regression" is a classification algorithm

$$y_\beta(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x + \varepsilon)}}$$

```
from sklearn.linear_model import LogisticRegression
#(just like LinearRegression)
```

```
from statsmodels.formula.api import Logit
#(just like OLS)
```