Classification

Andrey Sozykin
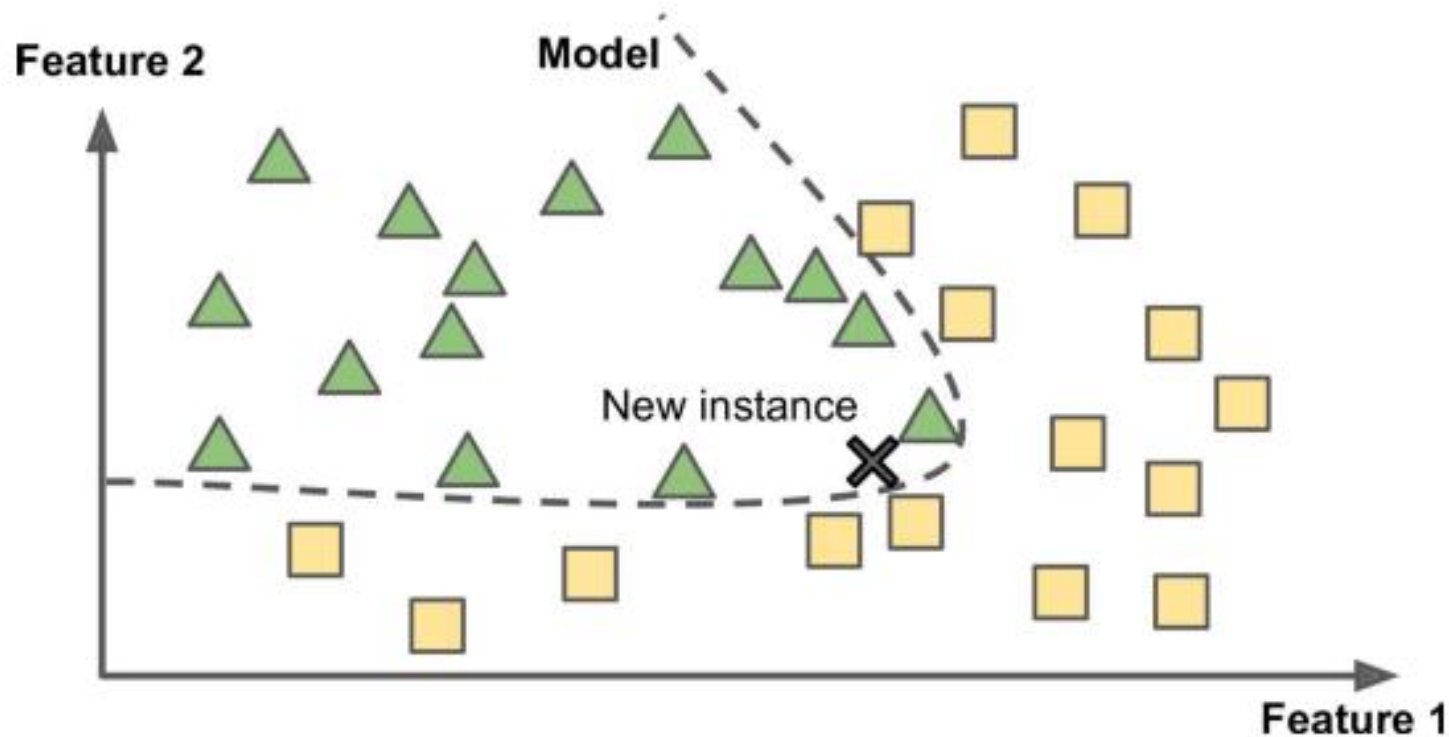
Andrey.Sozykin@urfu.ru

Regression

Source: finance.yahoo.com

Classification

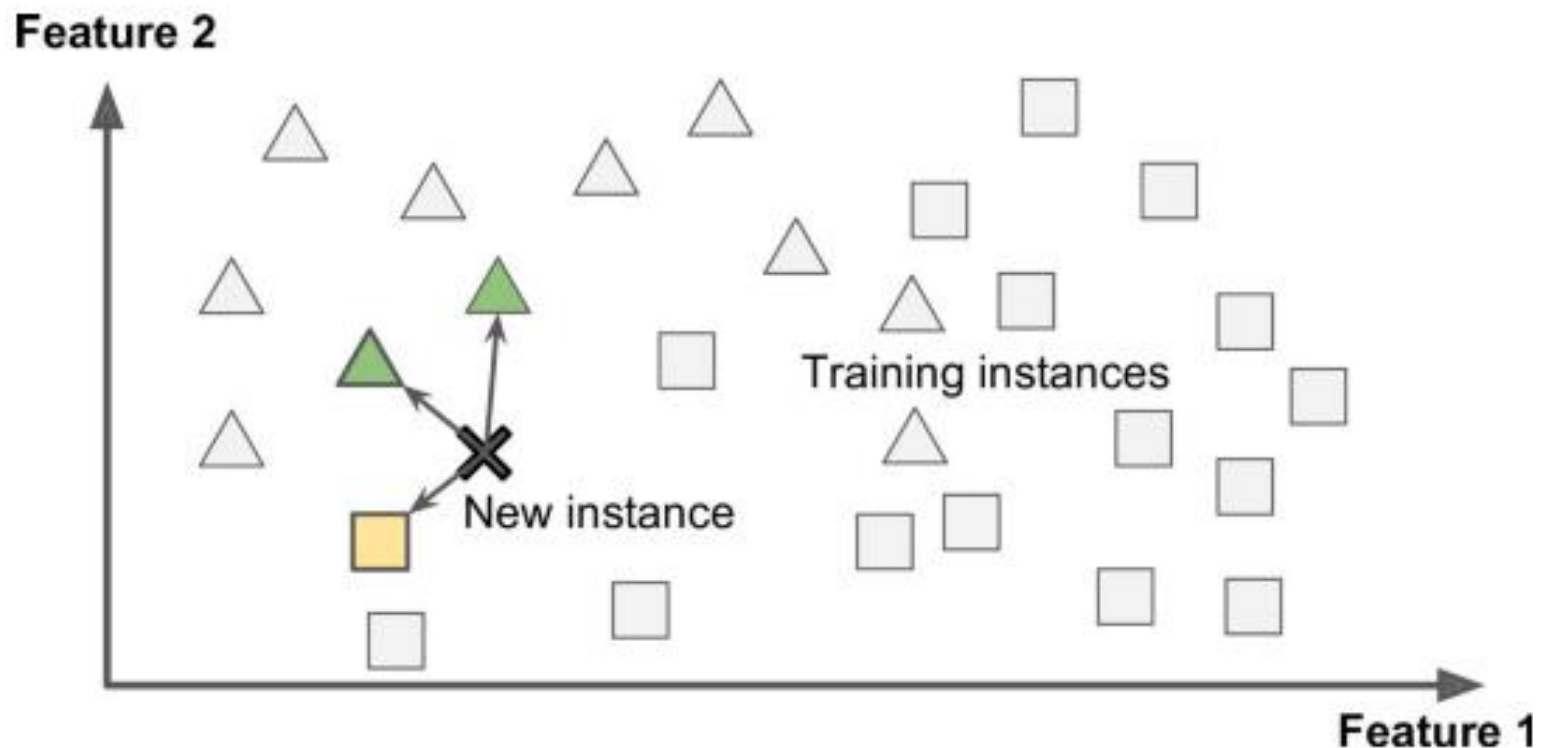# Model-based learning

Source: Géron, Aurélien. Hands-On Machine Learning with Scikit-Learn and TensorFlow

Instance-based learning

Feature 2

Training instances

✕ **New instance**

Feature 1

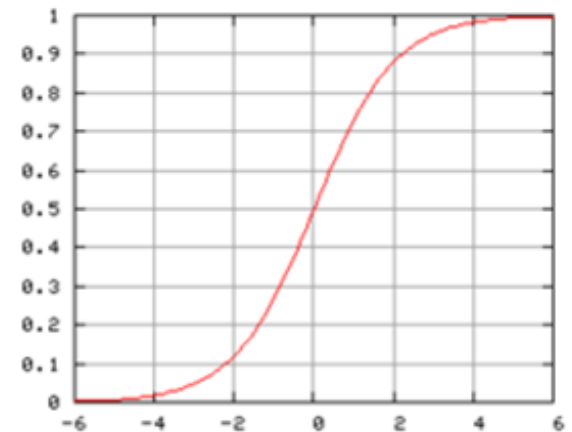Source: Géron, Aurélien. Hands-On Machine Learning with Scikit-Learn and TensorFlow

Logistic Regression

Linear Regression

$$y = w_0 \cdot 1 + w_1 \cdot x_1 + \cdots + w_n \cdot x_m$$
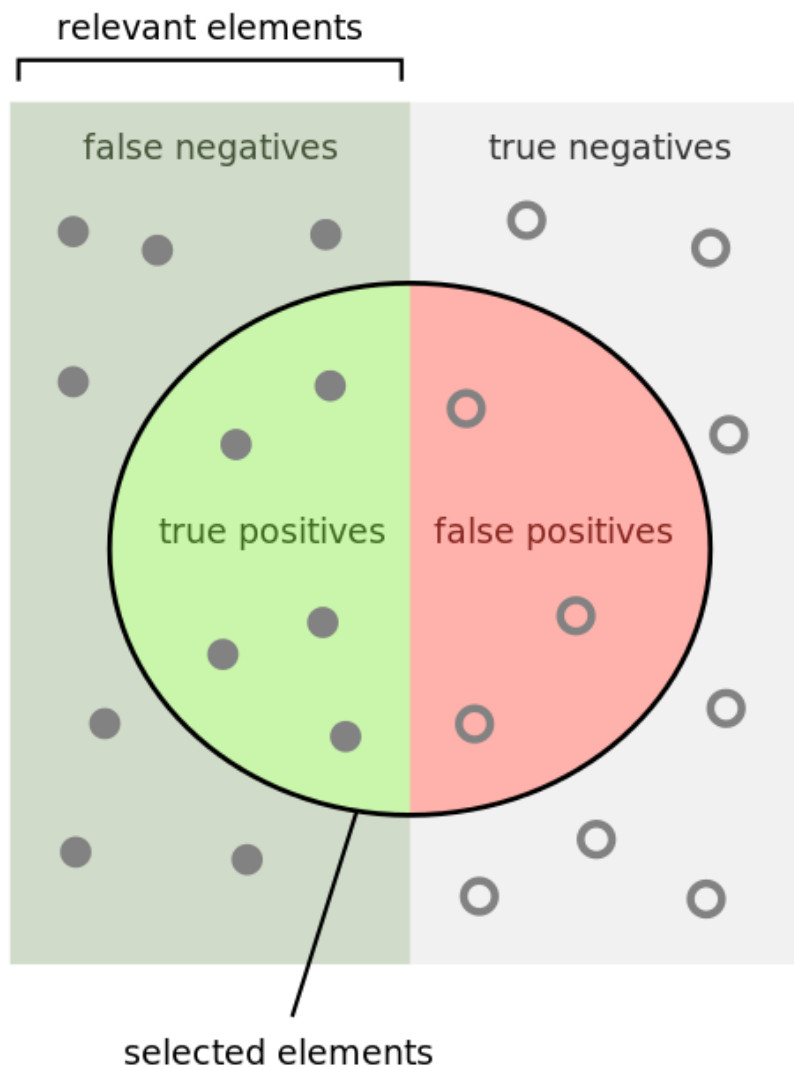
Logistic Regression

$$p = \frac{1}{1 + e^{-y}}$$

Accuracy

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies
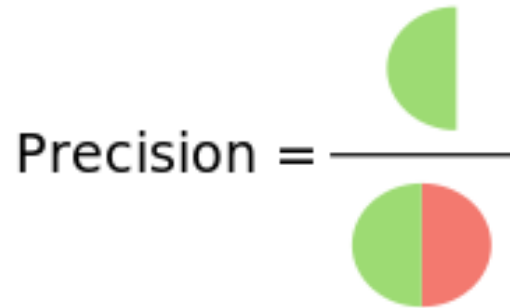
Accuracy is the fraction of predictions our model got right

$$accuracy = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ prediction}$$
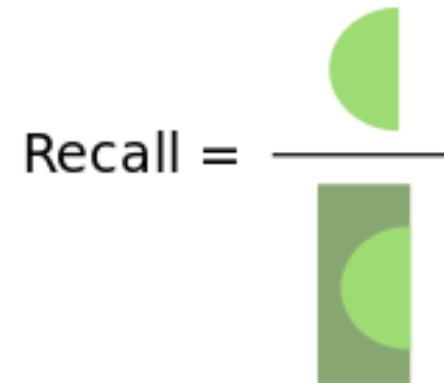
Confusion matrix

Precision and recall

How many selected items are relevant?

$$Precision = \frac{\text{(green half-circle)}}{\text{(green and red circle)}}$$

How many relevant items are selected?

$$Recall = \frac{\text{(green half-circle)}}{\text{(green box with half-circle)}}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

ROC/AUC

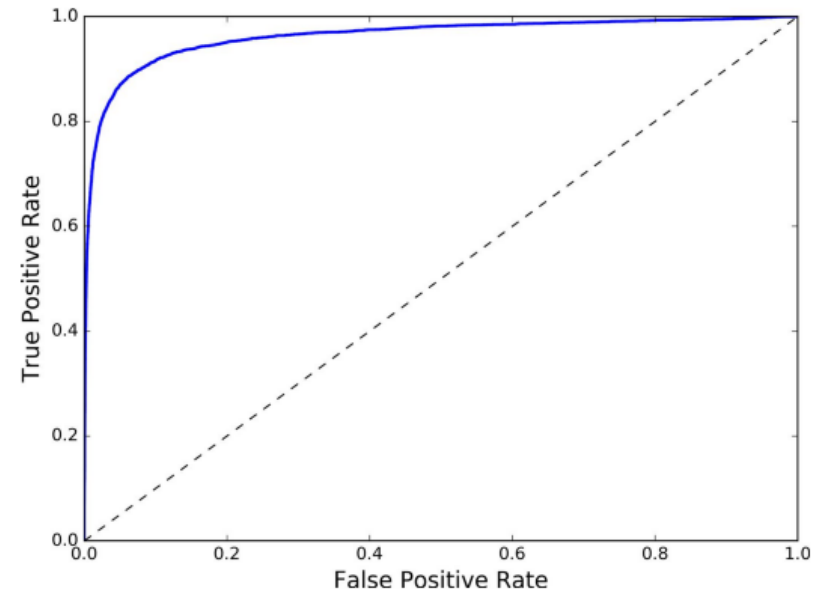Receiver Operating Characteristic – plot of true positive rate against the false positive rate.
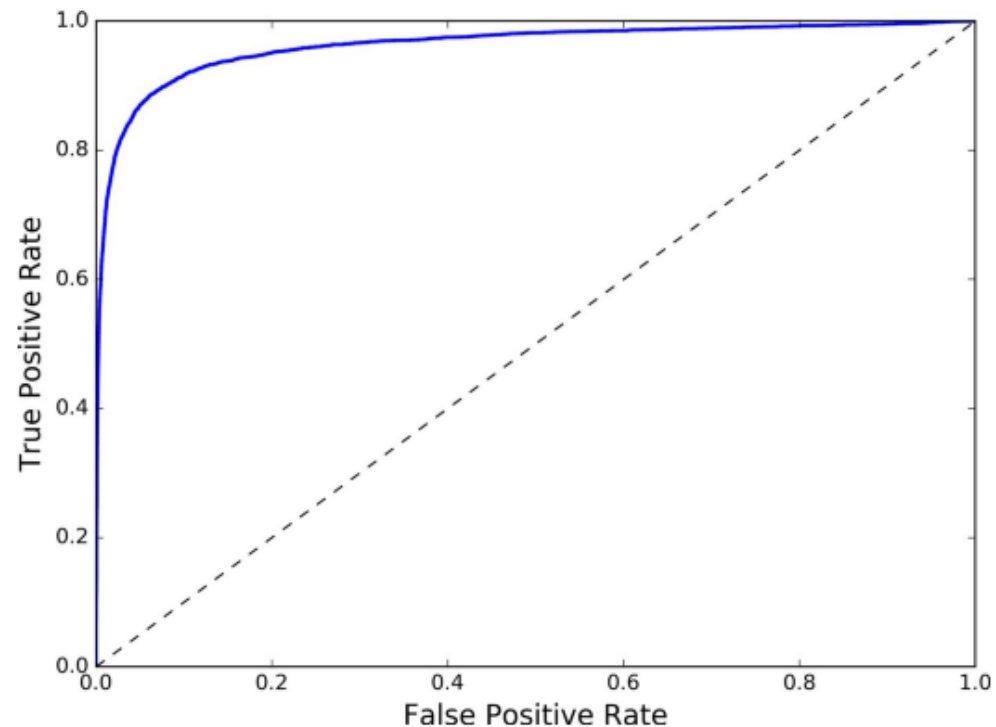
ROC/AUC

AUC – area under the curve

Good classifier – AUC close to 1

Bad classifier – AUC close to 0.5

Ural Federal University

named after the first President
of Russia B.N.Yeltsin

**Institute of radioelectronics
and information technologies**

Thank you!