أكاديمية طويق
Tuwaiq Academy

Introduction to Artificial Intelligence and Applications

Semester One | Course One

# Unit 3 : AI Concepts, Terminology and Application Domains PART II

# Key Concepts in ML

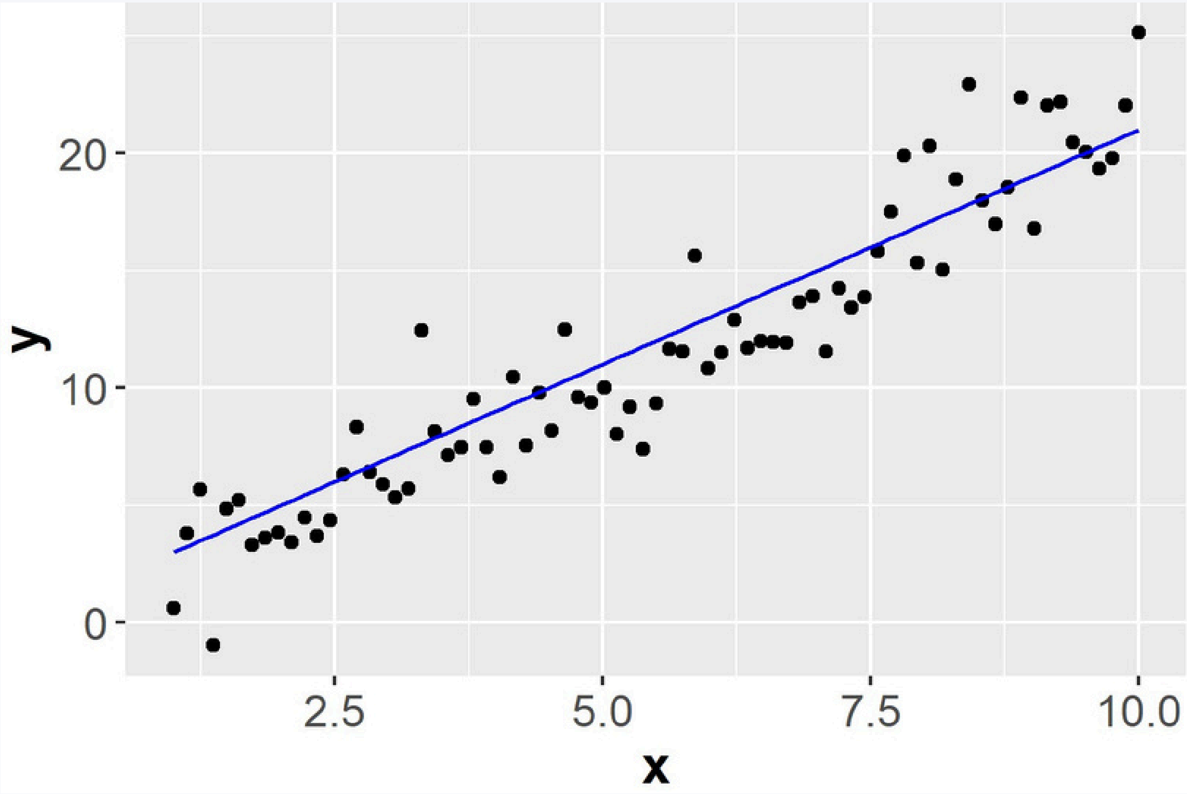
# SUPERVISED LEARNING

## REGRESSION VS CLASSIFICATION

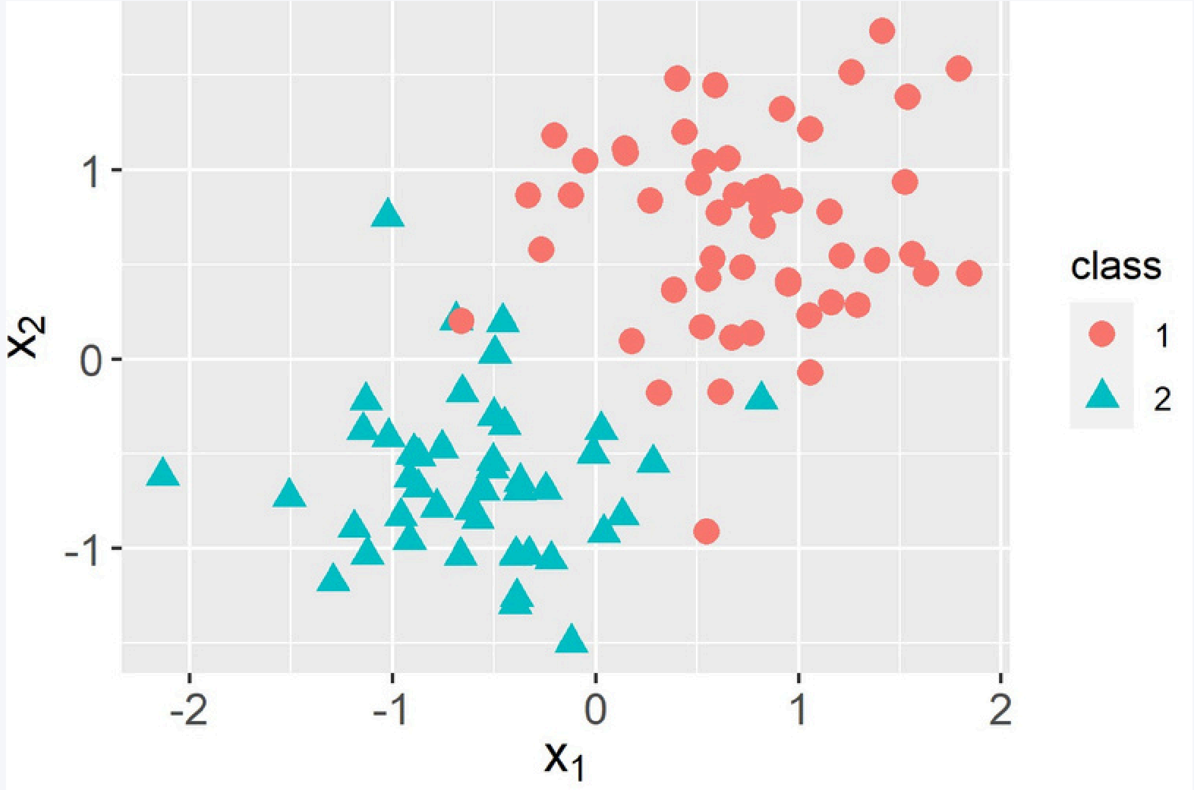
**Supervised tasks** are data situations where learning the functional relationship between inputs (features) and output (target) is useful.

The two most basic tasks are **regression and classification,** depending on whether the target is numerical or categorical.

**Regression**:Our observed labels come from **Y → R.**




**Classification**:Observations are categorized:**y ↑ Y={C1,...,Cg}.**




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# SUPERVISED LEARNING



## PREDICT VS. EXPLAIN

**We can distinguish two main reasons to learn this relationship:**

### Learning to Predict

- **Focus**: The outcome, not the structure or interpretability of the model.
- **Goal**: Make accurate predictions for new data.
- **Example**: Forecasting stock prices for practical decision-making.
- **Value**: Directly benefits by providing actionable predictions.

### Learning to Explain

- **Focus**: Gaining insights into the relationships within the data.
- **Goal**: Understand factors and their influence on outcomes.
- **Example**: Identifying risk factors for diseases.
- **Value**: Used to inform scientific, medical, or social discussions rather than operational predictions.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# SUPERVISED LEARNING

## REGRESSION EXAMPLE: HOUSE PRICES

### Predict the price for a house in a certain area

In this example, we might need to **learn to explain** so we can understand which features that influence the house price the most. But maybe we are also looking for underpriced houses and the predictor is of direct use, too.

| Features $x$ | | | | Target $y$ |
|---|---|---|---|---|
| square footage of the house | number of bedrooms | swimming pool (yes/no) | ... | house price in US$ |
| 1,180 | 3 | 0 | ... | 221,900 |
| 2,570 | 3 | 1 | ... | 538,000 |
| 770 | 2 | 0 | ... | 180,000 |
| 1,960 | 4 | 1 | ... | 604,000 |

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# SUPERVISED LEARNING

## REGRESSION EXAMPLE: LENGTH-OF-STAY

### Predict days a patient has to stay in hospital at time of admission

Learn to predict can be immensely **beneficial**, also it might be good for learning to explain.

| Features $x$ | | | | | Target $y$ |
|---|---|---|---|---|---|
| **diagnosis category** | **admission type** | **gender** | **age** | **...** | **Length-of-stay in the hospital in days** |
| heart disease | elective | male | 75 | ... | 4.6 |
| injury | emergency | male | 22 | ... | 2.6 |
| psychosis | newborn | female | 0 | ... | 8 |
| pneumonia | urgent | female | 67 | ... | 5.5 |

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# SUPERVISED LEARNING

## CLASSIFICATION EXAMPLE: RISK CATEGORY

**Predict one of five risk categories for a life insurance customer to determine the insurance premium**

Probably learn to predict, but the company might be **required to explain** its predictions to its customers.

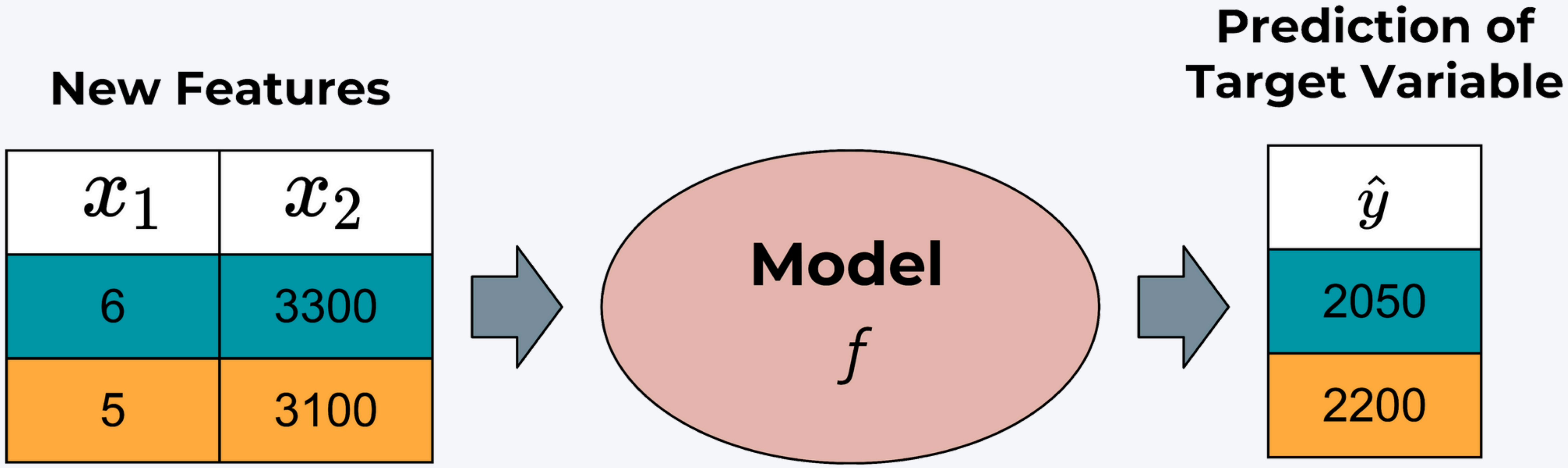
| Features $x$ | | | | Target $y$ |
|---|---|---|---|---|
| **job type** | **age** | **smoker** | **...** | **risk group** |
| carpenter | 34 | 1 | ... | 3 |
| stuntman | 25 | 0 | ... | 5 |
| student | 23 | 0 | ... | 1 |
| white-collar worker | 39 | 0 | ... | 2 |

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# MODELS & PARAMETERS

A **model** (or hypothesis) **f:X→Rg** is a function that maps feature vectors to predicted target values.

- **Regression**: **g=1**, where the output is a single predicted value.
- **Classification**: **g** is the number of classes, and the output consists of scores or class probabilities.




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

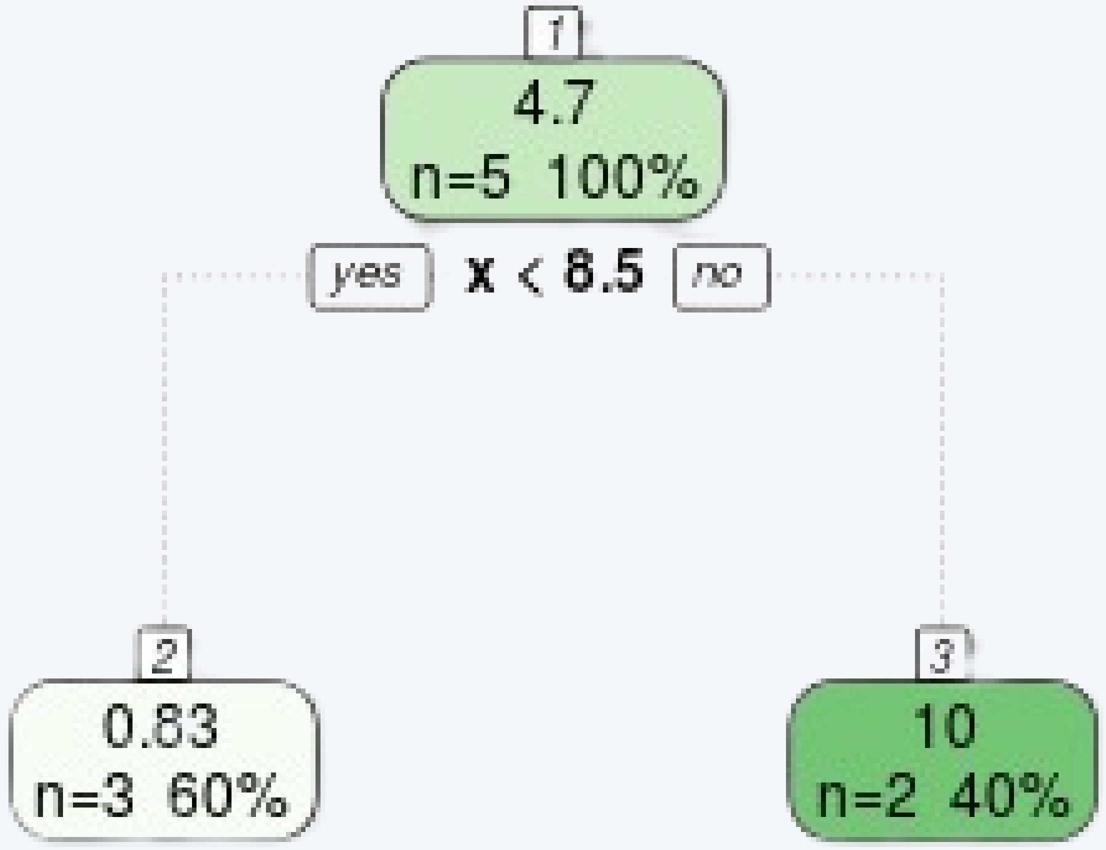
# MODELS & PARAMETERS

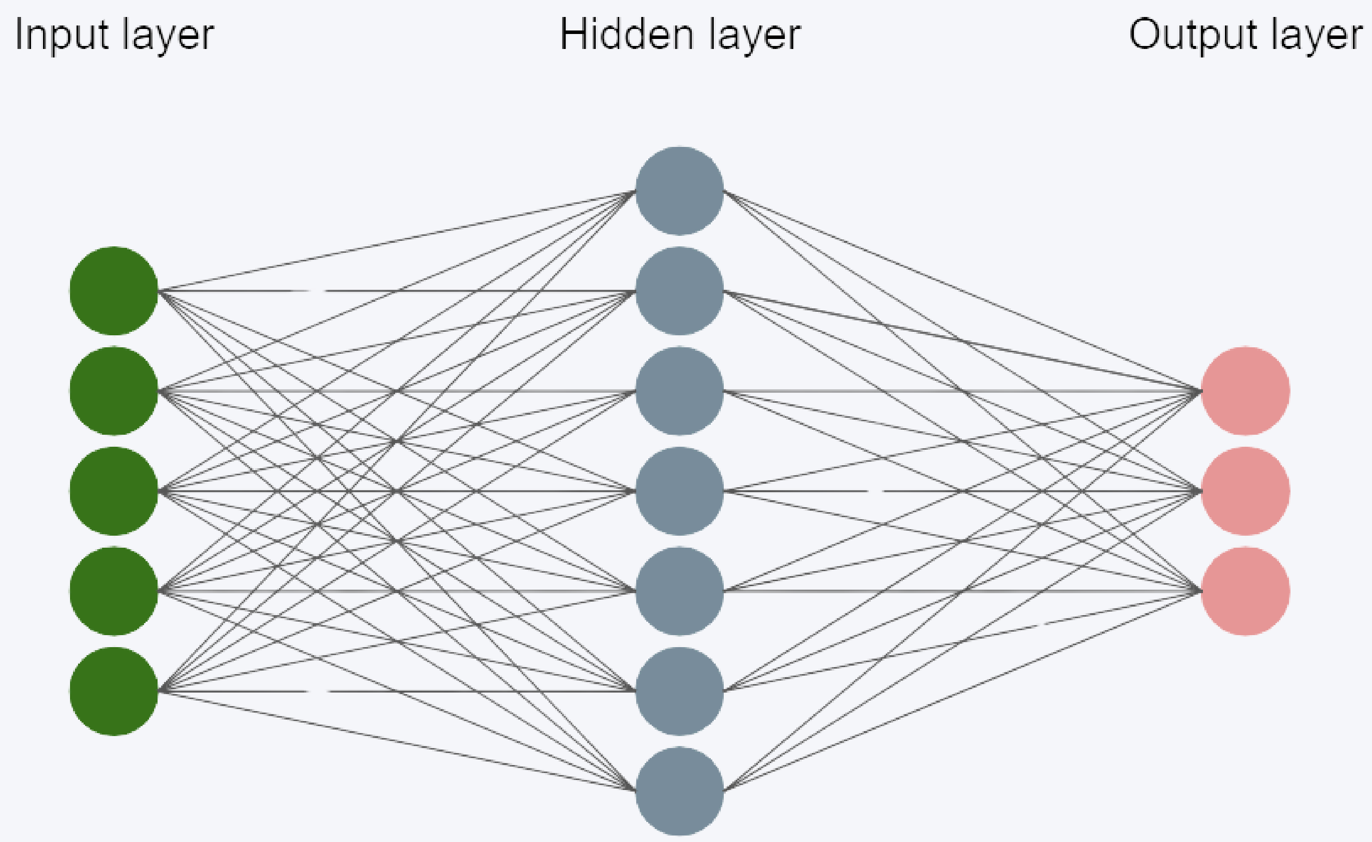
أكاديمية طويق
Tuwaiq Academy

## WHAT IS A MODEL?

The function **f** is designed to capture intrinsic patterns in the data, with the assumption that these patterns hold for all data drawn from **Pxy.**

- Models can vary in **complexity**, from simple (e.g., linear models, tree stumps) to complex (e.g., deep neural networks).
- There are infinite ways to construct such functions, depending on the task and data.




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# MODELS & PARAMETERS

أكاديمية طويق
Tuwaiq Academy

## HYPOTHESIS SPACES

- Think of a hypothesis space as a menu of all possible models that we can choose from for a specific problem.
- Each model on the menu has a specific way of making predictions.
- Without narrowing down the options, finding the "**best model**" would be impossible because the choices would be unlimited.
- **Key Idea:** To make this manageable, we restrict the menu to a specific type of models (like linear models, neural networks, etc.), which is called a structural prior

## PARAMETRIZATION

- Imagine each model in the hypothesis space as a recipe, and the ingredients (parameters) in the recipe determine the specific dish (model).
- **Parametrization** means using these ingredients (parameters) to adjust and fine-tune the recipe to make the model work for the given data.
- All the models in the hypothesis space share the same basic structure (e.g., all are linear equations or neural networks).
- The specific differences between the models come from the parameters we use to define them.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# MODELS & PARAMETERS

## HYPOTHESIS SPACES

- **Parameters** are like knobs or dials you can adjust to make the model perform better.

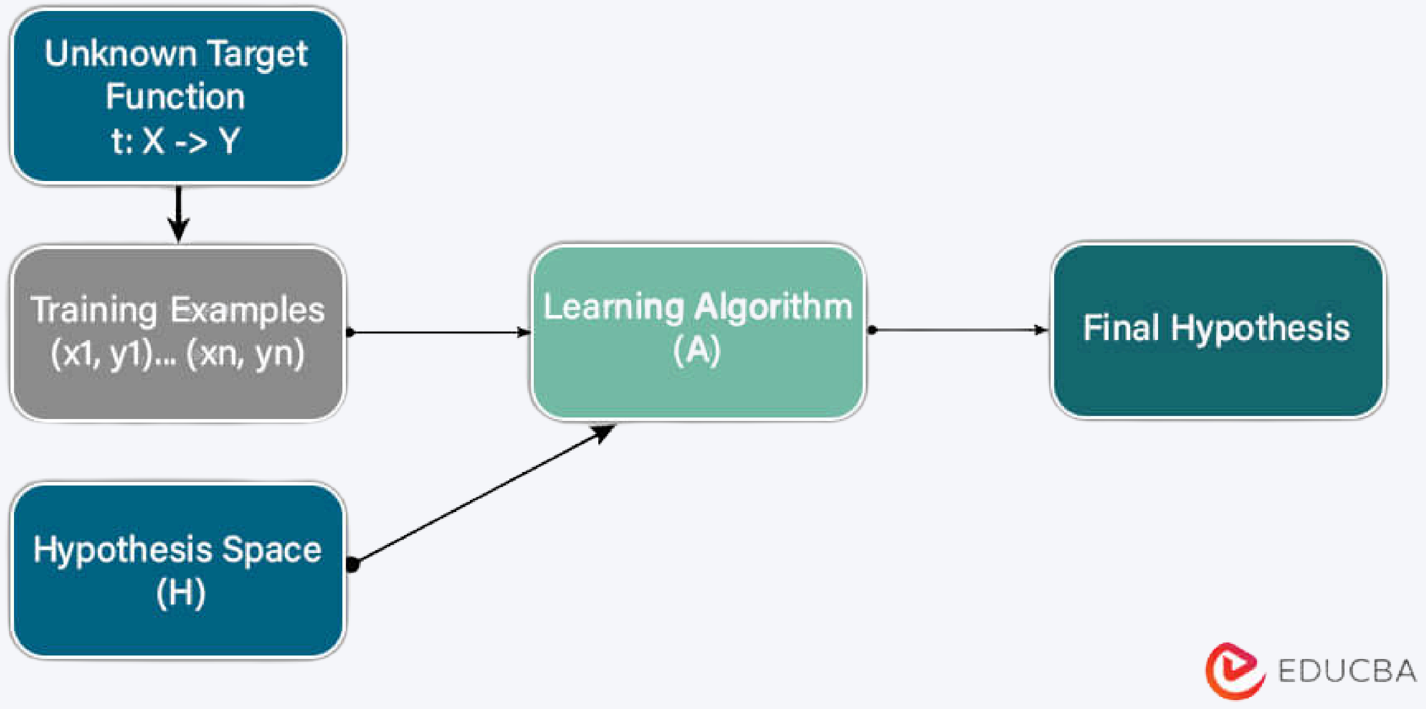
- In a straight-line model (like predicting trends), the slope and intercept are parameters that decide how steep or flat the line is.

- When we "**set**" these parameters, the model becomes fully defined and ready to make predictions.

Instead of thinking of a **hypothesis space** as just a menu of models, think of it as a **collection of all possible ways you can tweak the knobs (parameters) to get the best model for your data.**

- The whole point of **parametrization** is to help us find the best possible model for our data within the hypothesis space by tuning the parameters.

- This process is what makes machine learning models adaptable to different problems!




https://www.educba.com/hypothesis-in-machine-learning/

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# LEARNER

## TRAINING PROCEDURE OF ML SUPERVISED MODEL:

### Understand that a supervised learner fits models automatically from training data

- Imagine we want to investigate how working conditions affect productivity of employees.
- It is a **regression task** since the target productivity is continuous.

**Data Collected:** Worked minutes per week (productivity), how many people work in the same office as the employee in question, and the employee's salary.

| Features $x$ | | Target $y$ |
|---|---|---|
| People in Office (Feature 1) $x_1$ | Salary (Feature 2) $x_2$ | Worked Minutes Week (Target Variable) |
| 4 | 4300 € | 2220 |
| 12 | 2700 € | 1800 |
| 5 | 3100 € | 1920 |

$n = 3$

$p = 2$

$x_1^{(2)}$ $x_2^{(1)}$ $y^{(3)}$

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# LEARNER

## TRAINING PROCEDURE OF ML SUPERVISED MODEL:

**Automatically identify the fundamental functional relation in the data that maps an object's features to the target.**

- **Supervised learning** means we make use of labeled data for which we observed the outcome.
- We use the labeled data to learn a model **f.**
- Ultimately, we use our model to compute predictions for new data whose target values are unknown.

| Features $x$ | | Target $y$ |
|---|---|---|
| People in Office (Feature 1) $x_1$ | Salary (Feature 2) $x_2$ | Worked Minutes Week (Target Variable) |
| 4 | 4300 € | 2220 |
| 12 | 2700 € | 1800 |
| 5 | 3100 € | 1920 |

$n = 3$

$p = 2$

$x_1^{(2)}$ $x_2^{(1)}$ $y^{(3)}$

| | $x_1$ | $x_2$ | Functional Relationship | $y$ |
|---|---|---|---|---|
| Already seen Data | 4 | 4300 € | $f$ | 2200 |
| | 12 | 2700 € | | 1800 |
| | 15 | 3100 € | | 1920 |
| New Data | 6 | 3300 € | | ??? |
| | 5 | 3100 € | | ??? |

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# LEARNER

## WHAT IS LEARNER ? !

**A learner (or learning algorithm/inducer) is the method used to find the best model (f) for a given dataset.**

**Role**: Selects the best function from a prescribed hypothesis space (H) based on the training data.

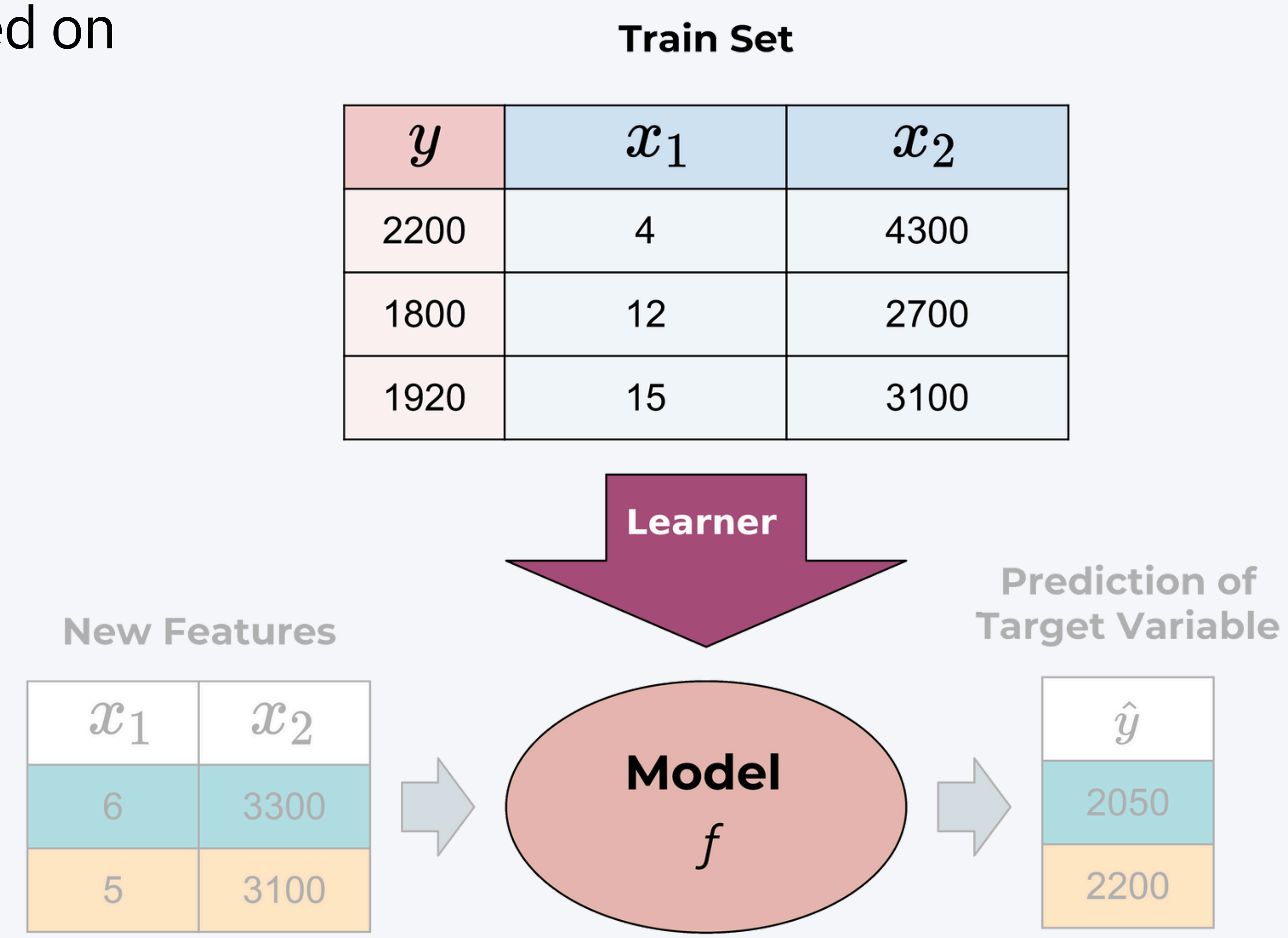
**Input:**

- Training data (D): The dataset used to train the model.
- Hyperparameters (ω): Settings that control how the model is built or optimized.

**Output:** A specific model that best fits the data.

**Mapping Concept:**

- Learner: **D+ω → f∈H**
- In simple terms, the learner takes the data and hyperparameter settings and uses them to choose the best-fitting model.




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# LOSS IN ML

## HOW TO EVALUATE MODELS?

- When training a **learner**, we **optimize** over our hypothesis space, to find the function which matches our training data best.
    - This means, we are looking for a **function**, where the predicted output per training point is **as close as possible** to the observed label.

| Features $x$ | | Target $y$ | | Prediction $\hat{y}$ |
|---|---|---|---|---|
| People in Office (Feature 1) $x_1$ | Salary (Feature 2) $x_2$ | Worked Minutes Week (Target Variable) | | Worked Minutes Week (Target Variable) |
| 4 | 4300 € | 2220 | $\overset{?}{\approx}$ | 2588 |
| 12 | 2700 € | 1800 | | 1644 |
| 5 | 3100 € | 1920 | | 1870 |

$\mathcal{D}_{\text{train}}$

To make this **precise**, we need to define now how we **measure** the **difference** between a prediction and a ground truth label pointwise.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

أكاديمية طويق
Tuwaiq Academy

# LOSS IN ML

## WHAT IS LOSS?

A **loss** function measures *how far off a model's prediction is from the true value for a single observation*.

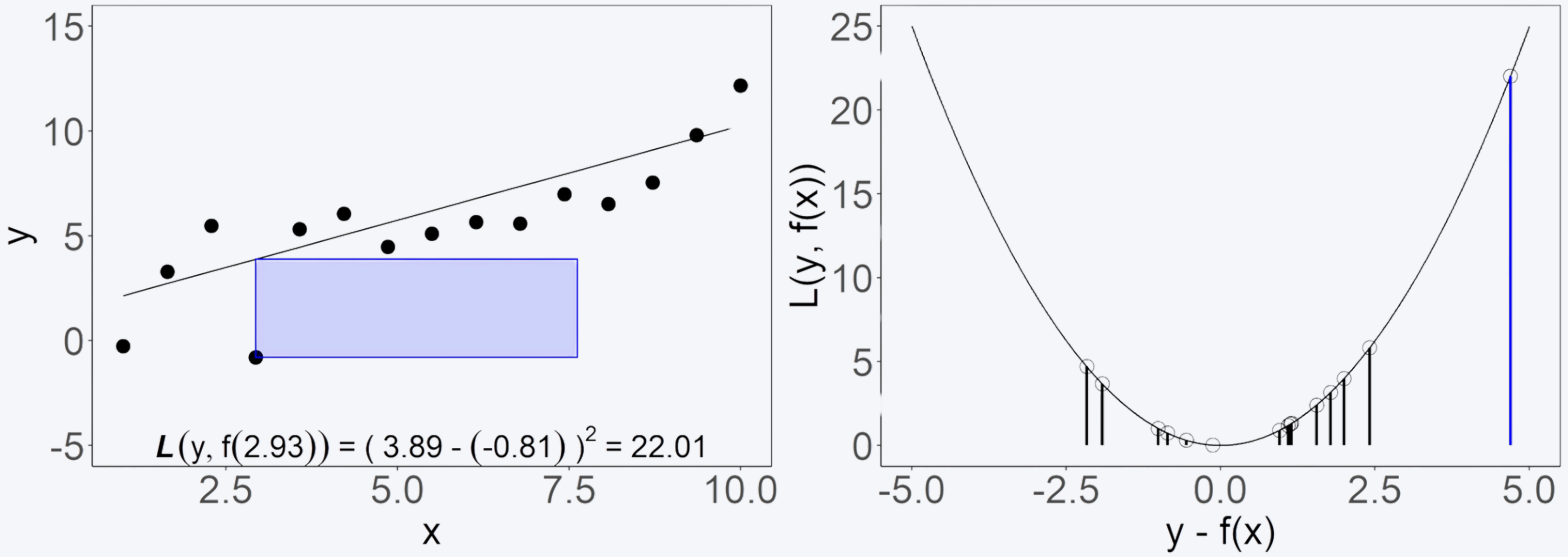
It quantifies the "**quality**" of the prediction.

**Role:**

- Smaller loss means better predictions; larger loss indicates poor predictions.

- Types of Loss Functions:
    - L2 Loss (Squared Error): $$L(y, f(x)) = (y - f(x))^2$$

- Penalizes larger errors more heavily.
- Common in regression tasks.




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

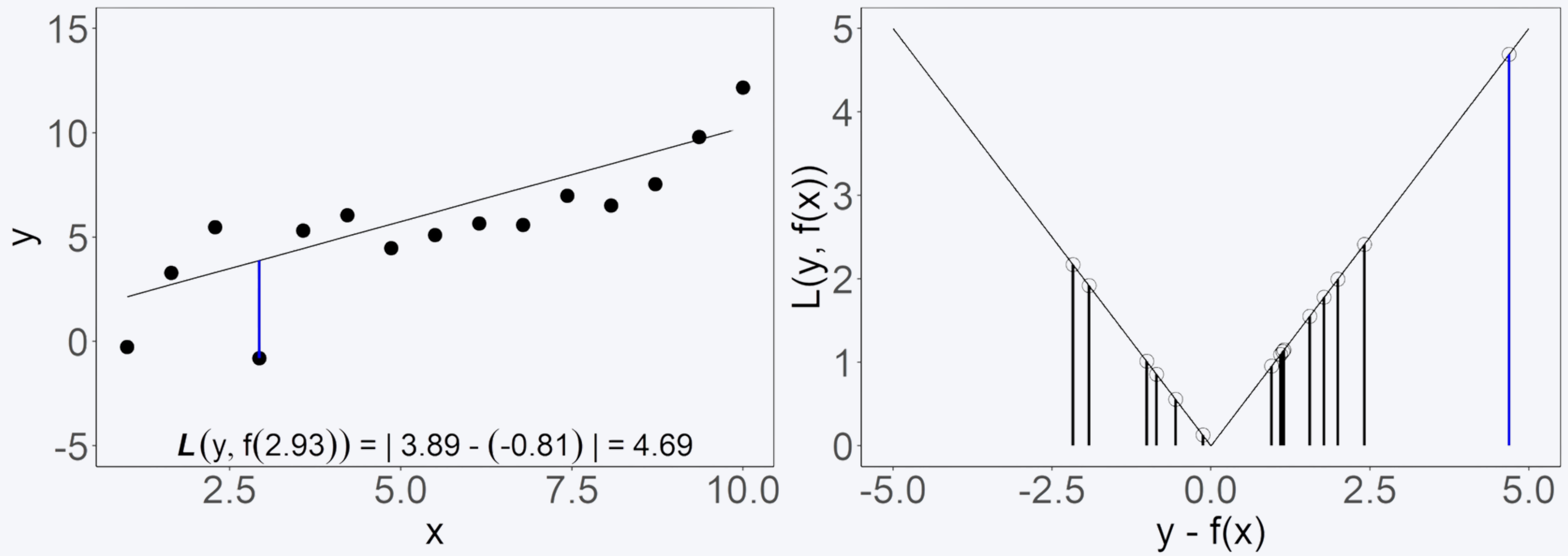
# LOSS IN ML

أكاديمية طويق
Tuwaiq Academy

## TYPES OF LOSS

A **loss** function measures *how far off a model’s prediction is from the true value for a single observation*.

- **Absolute Loss:** $$L(y, f(x)) = |y - f(x)|$$
  - Treats all errors equally, regardless of size.




- **Mapping:**
- **The loss function takes:**
  - True value (y)
  - Predicted value f(x)
  - and computes a numerical value indicating error.

In simple terms, the loss function tells us how "**wrong**" the model's predictions are for each observation.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# OPTIMIZATION IN ML

Tuwaiq Academy أكاديمية طويق

## WHAT DO WE MEAN BY OPTIMIZATION ?

**Optimization** is the process of finding the **best** model parameters that **minimize** a given loss function (i.e., reduce prediction errors).

It ensures the **model learns patterns** in the data effectively.

## Gradient Descent

- **Gradient Descent** is an **iterative** algorithm used to **minimize** the loss function by updating model parameters step by step.
    a. Calculate the gradient (slope) of the loss function with respect to the parameters.
    b. Update parameters in the opposite direction of the gradient to reduce the loss.

- **Formula:** $\theta \leftarrow \theta - \eta \cdot \nabla L(\theta)$

    - $\theta$: Parameters to update.
    - $\eta$: Learning rate (step size).
    - $\nabla L(\theta)$: Gradient of the loss function.

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# OPTIMIZATION IN ML

أكاديمية طويق
Tuwaiq Academy

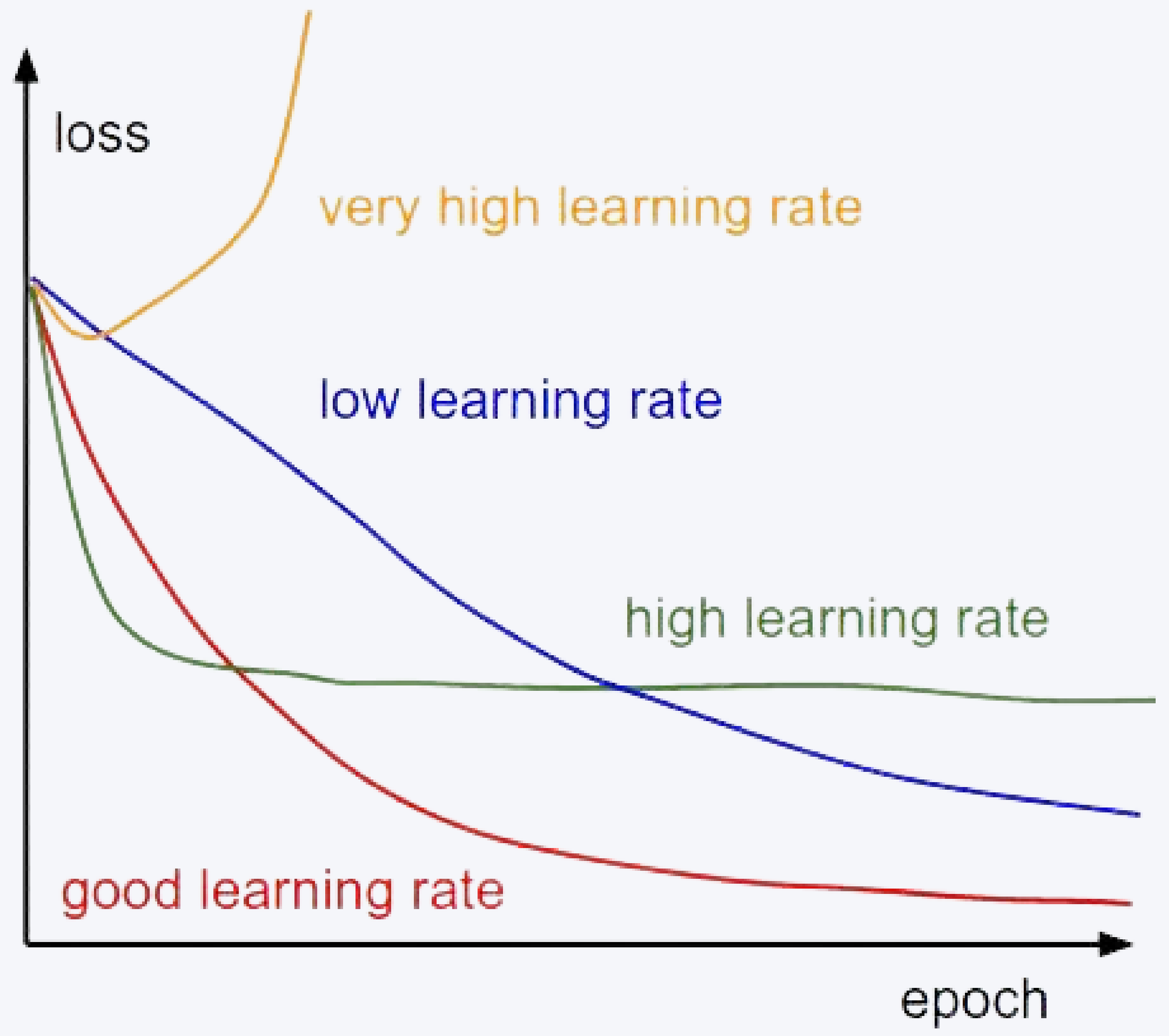
## WHAT DO WE MEAN BY OPTIMIZATION ?

### Learning Rate:

The learning rate (η) is the step size used to update model parameters during optimization.

- ○ **Too small**: Slow learning (takes many iterations).
- ○ **Too large:** Can overshoot the minimum or cause instability.

- Finding the balance is crucial for efficient training.




https://towardsdatascience.com/https-medium-com-dashingaditya-rakhecha-understanding-learning-rate-dd5da26bb6de

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# REGRESSION IN ML



### Linear Regression

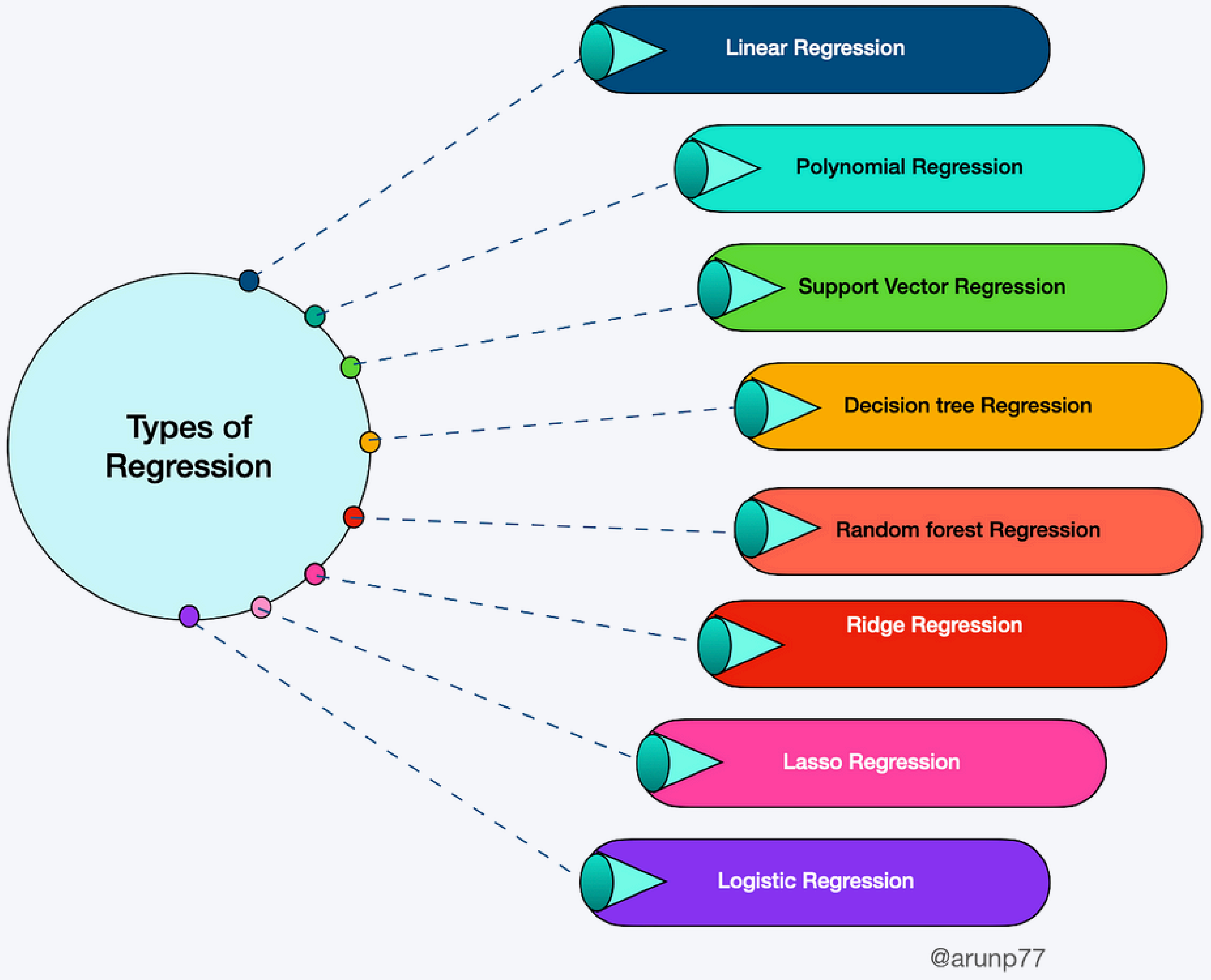
- **Definition**: Predicts a continuous outcome by finding the straight-line relationship between input variables (features) and the output.
- **Where to Use:** When the relationship between input and output is roughly linear.
- **Example**: Predicting house prices based on size, location, and number of rooms.

### Polynomial Regression

- **Definition**: Fits a curve to the data, rather than a straight line, to capture non-linear relationships.
- **Where to Use:** When the data shows a curved or complex trend instead of a straight-line relationship.
- **Example**: Predicting the growth of plants over time when the growth rate slows down after an initial burst.

### Logistic Regression

- **Definition**: Used to predict probabilities and classify data into categories (e.g., yes/no, true/false).
- **Where to Use:** For binary or multi-class classification problems.
- **Example**: Predicting whether an email is spam (yes or no).




https://arunp77.medium.com/regression-algorithms-29f112797724

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# REGRESSION IN ML

أكاديمية طويق
Tuwaiq Academy

## Ridge Regression (L2 Regularization)

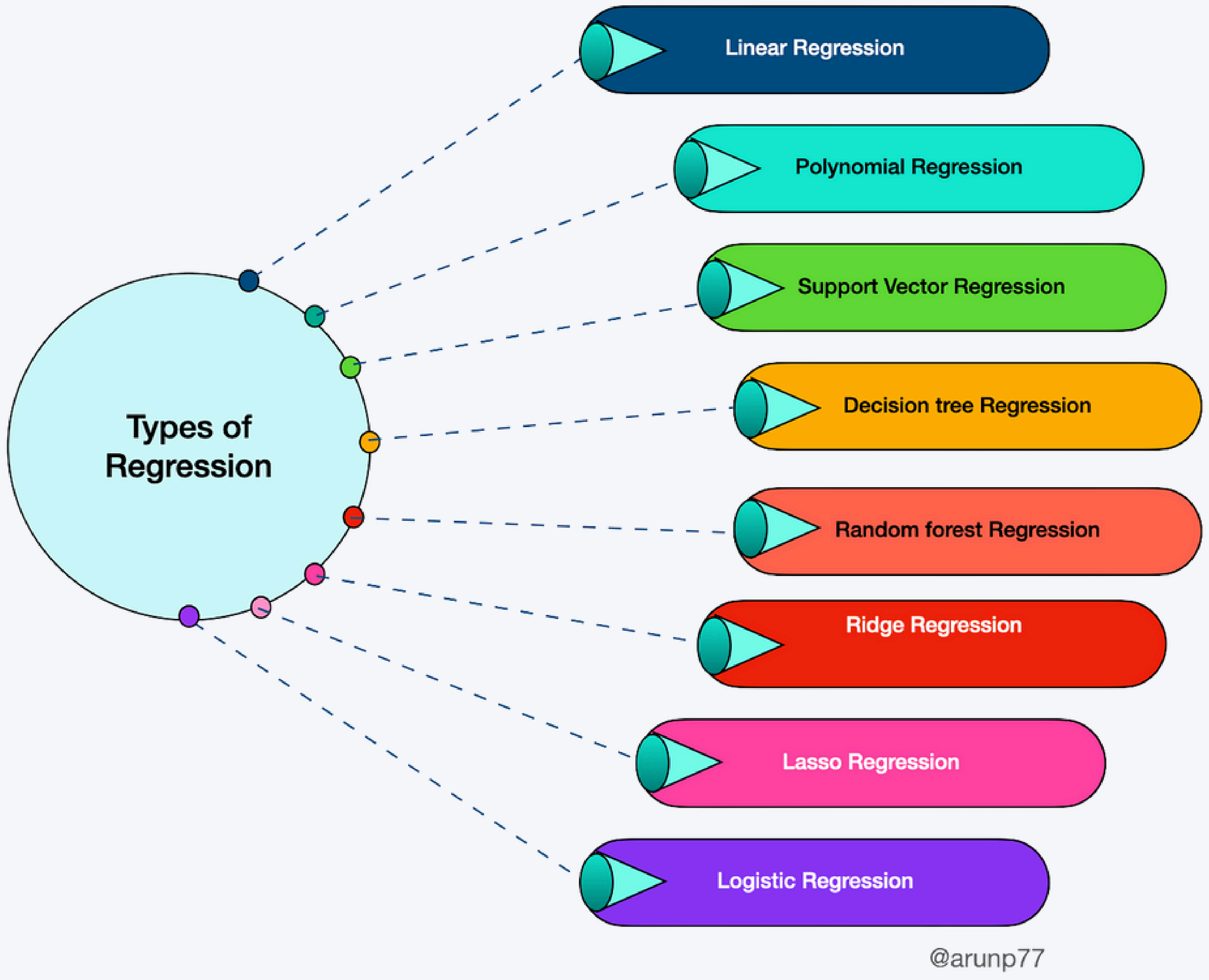
- **Definition**: Adds a penalty to prevent the model from being too sensitive to small changes in the data.
- **Where to Use:** When the data has many correlated variables, and you want to avoid overfitting.
- **Example**: Predicting demand for products with many related factors like price, weather, and season.

## Lasso Regression (L1 Regularization)

- **Definition**: Similar to Ridge Regression but also shrinks some variables to zero, effectively selecting only the most important features.
- **Where to Use:** For feature selection in high-dimensional data.
- **Example**: Identifying the most important factors in predicting a patient's risk of heart disease from a large set of health indicators.

## Support Vector Regression (SVR)

- **Definition**: Finds the best-fit line or curve within a certain margin of tolerance to capture the data's trends.
- **Where to Use:** For complex, non-linear relationships with high-dimensional data.
- **Example**: Predicting stock prices based on a variety of technical indicators.




https://arunp77.medium.com/regression-algorithms-29f112797724

Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

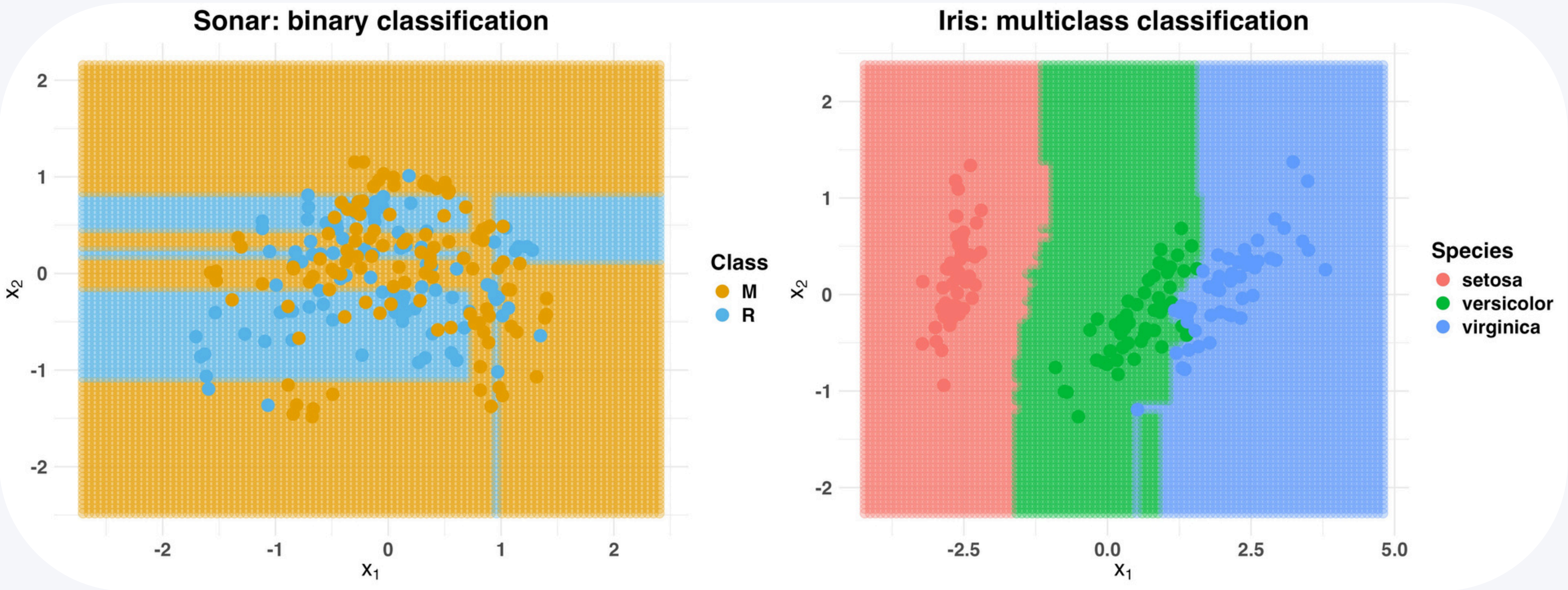
# CLASSIFICATION IN ML

## WHAT IS CLASSIFICATION ?

**Learn functions that assign class labels to observation / feature vectors. Each observation belongs to exactly one class. The main difference to regression is that the target is categorical.**

## BINARY AND MULTICLASS TASKS

- **Tasks have a finite number of (unordered) classes. They can be binary or multiclass.**




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

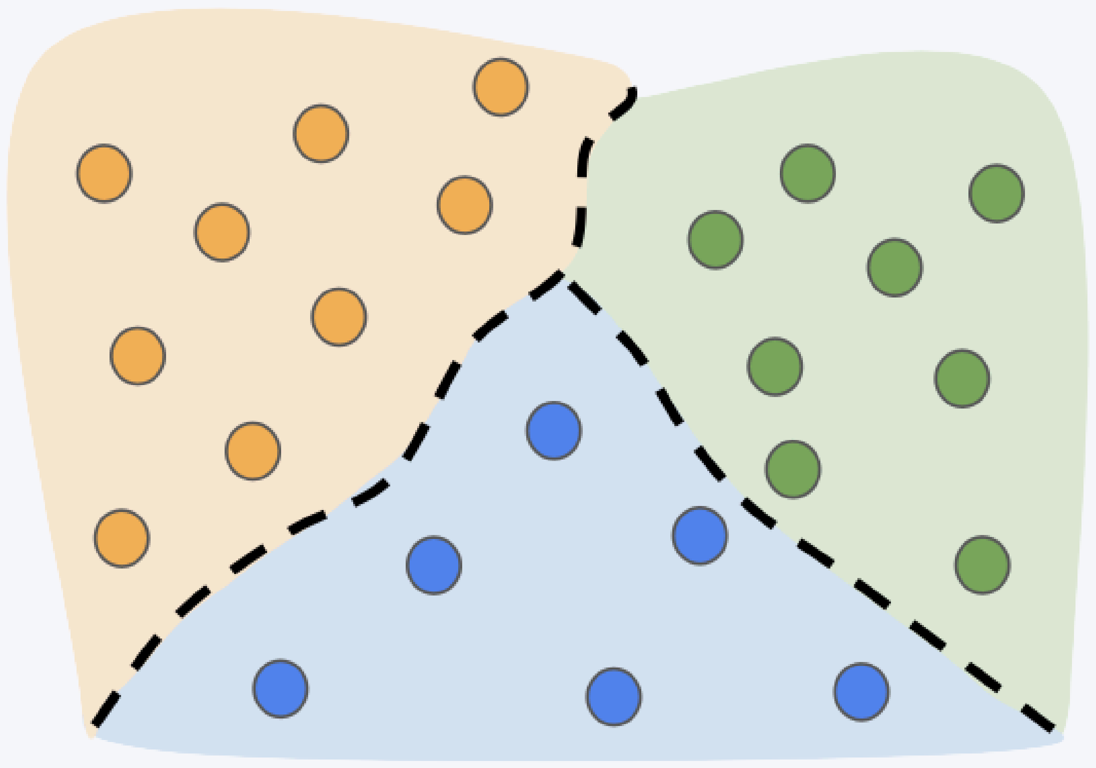
# CLASSIFICATION IN ML

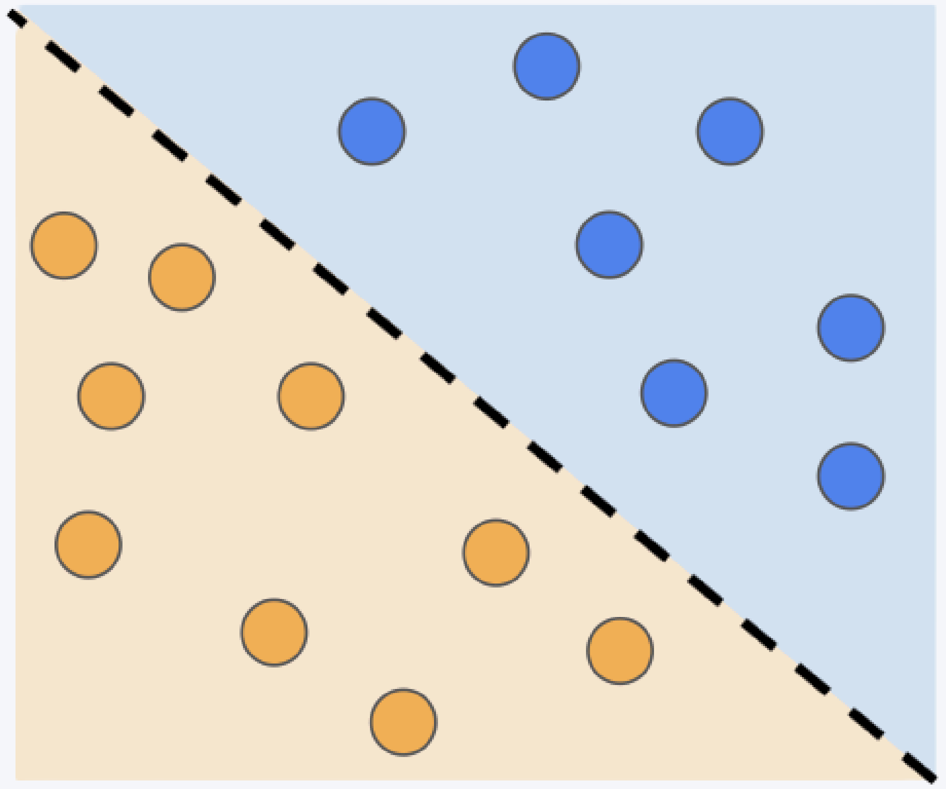
أكاديمية طويق
Tuwaiq Academy

## What are Decision Boundaries?

- A decision boundary is a line (or surface in higher dimensions) that separates different classes in the feature space.
- It helps the model decide which class a data point belongs to based on its features.

In machine learning, decision boundaries can be straight lines (linear models) or curved/complex shapes (non-linear models) depending on the algorithm used!





Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.

# CLASSIFICATION IN ML

أكاديمية طويق
Tuwaiq Academy

## Logistic Regression

- **Decision Boundary**: A straight line, as logistic regression is a linear classifier.
- **Behavior**: Works well when the data is linearly separable (can be split using a straight line).

## Naive Bayes

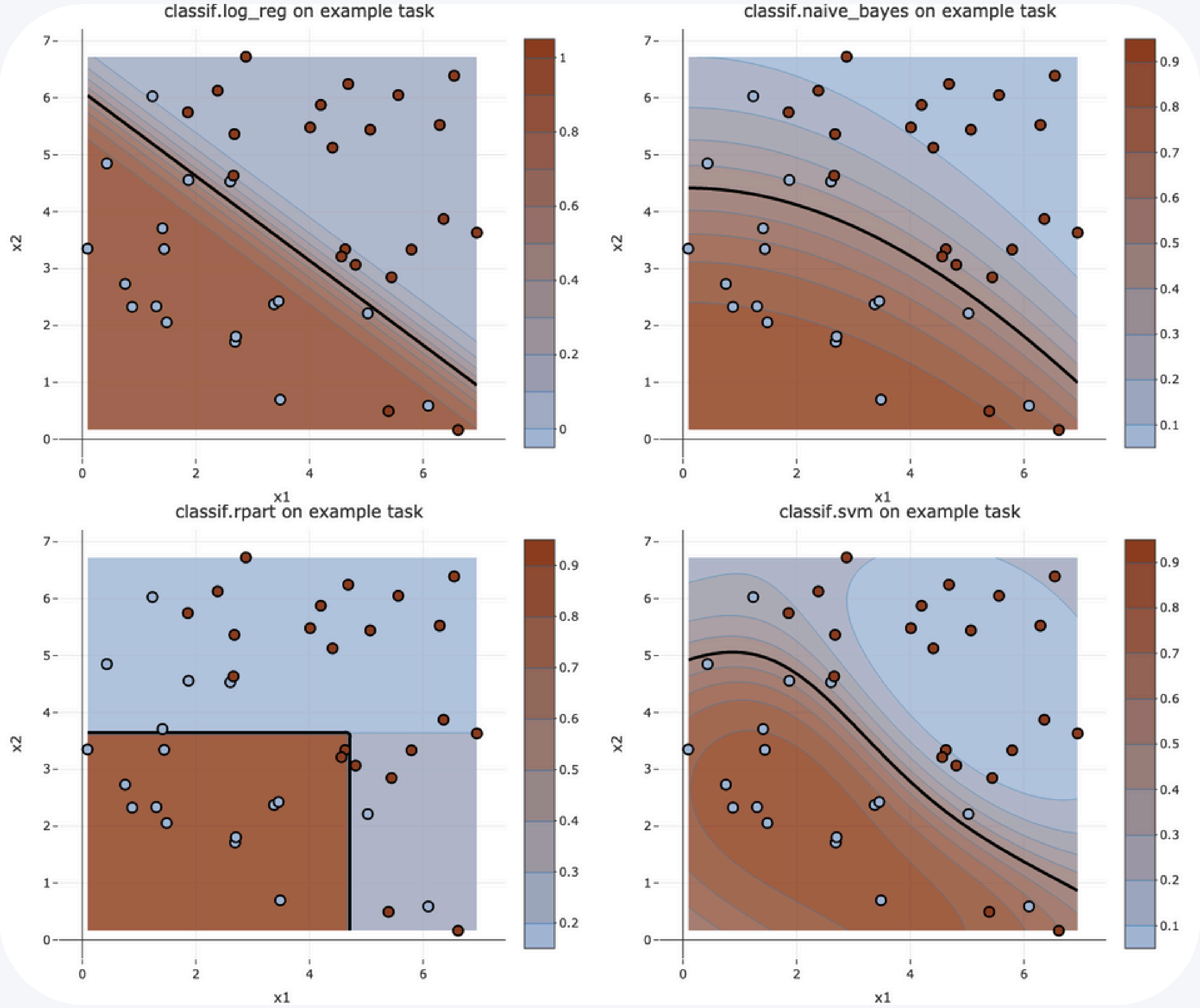
- **Decision Boundary:** Curved lines, showing how Naive Bayes considers probabilistic assumptions about the data.
- **Behavior**: Handles more complex relationships compared to logistic regression but may not perfectly separate classes.

## Decision Tree (rpart)

- **Decision Boundary:** Step-like, creating rectangular regions based on feature splits.
- **Behavior**: Divides the space using simple rules, useful for interpretable models, but can miss smooth transitions between classes.

## Support Vector Machine (SVM)

- **Decision Boundary:** Smooth, non-linear curves, showing SVM's ability to handle more complex patterns.
- **Behavior**: Creates flexible boundaries using kernels, great for datasets with intricate class separations




Bothmann, L., Strickroth, S., Casalicchio, G., Rügamer, D., Lindauer, M., Scheipl, F. & Bischl, B.. (2023). Developing Open Source Educational Resources for Machine Learning and Data Science. <i>Proceedings of the Third Teaching Machine Learning and Artificial Intelligence Workshop</i>, in <i>Proceedings of Machine Learning Research</i> 207:1-6 Available from https://proceedings.mlr.press/v207/bothmann23a.html.