Support Vector Machine

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

What is SVR ?

- The **SVM (Support Vector Machines)** is a **supervised learning model** used for classification or regression analysis.

- In **regression** problems the SVM is called **SVR (Support Vector Regression)**, and it is used to estimate the relationships between a dependent variable (often called the outcome variable) and one or more independent variables (often called predictors, covariates, or features) .

Examples:

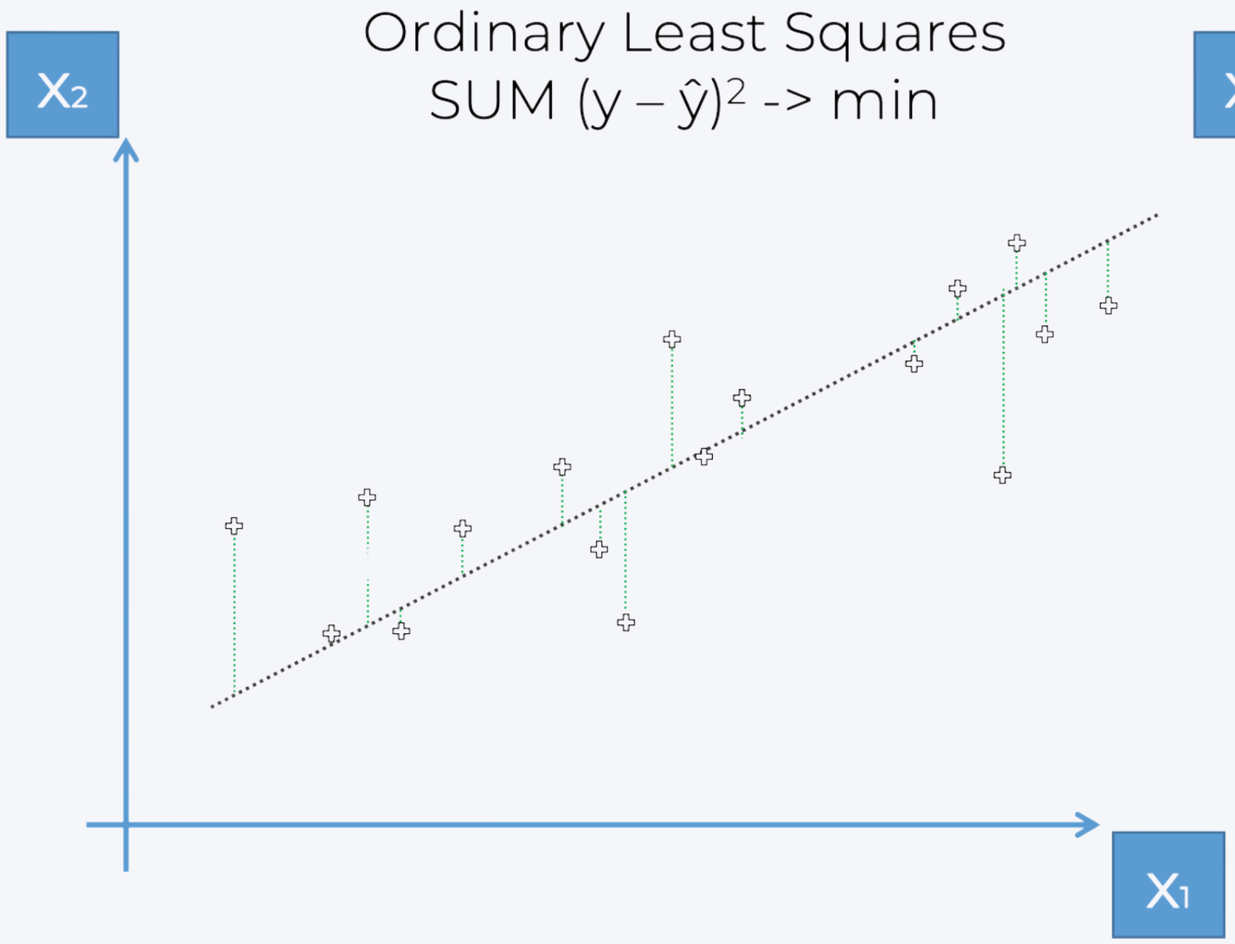
- Predicting disease progression.
- Engineering data (e.g., stress-strain curves).

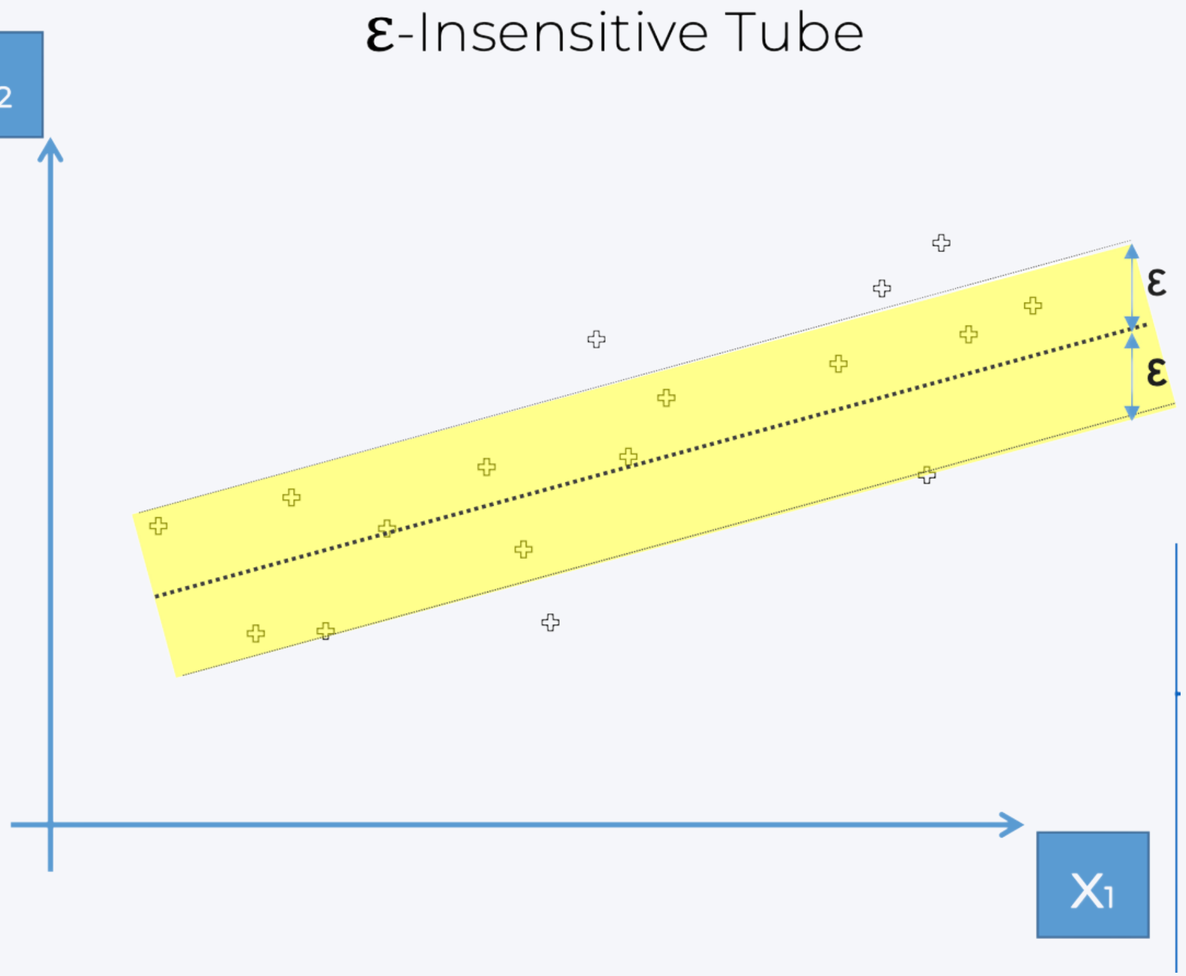
https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

What is SVR ?



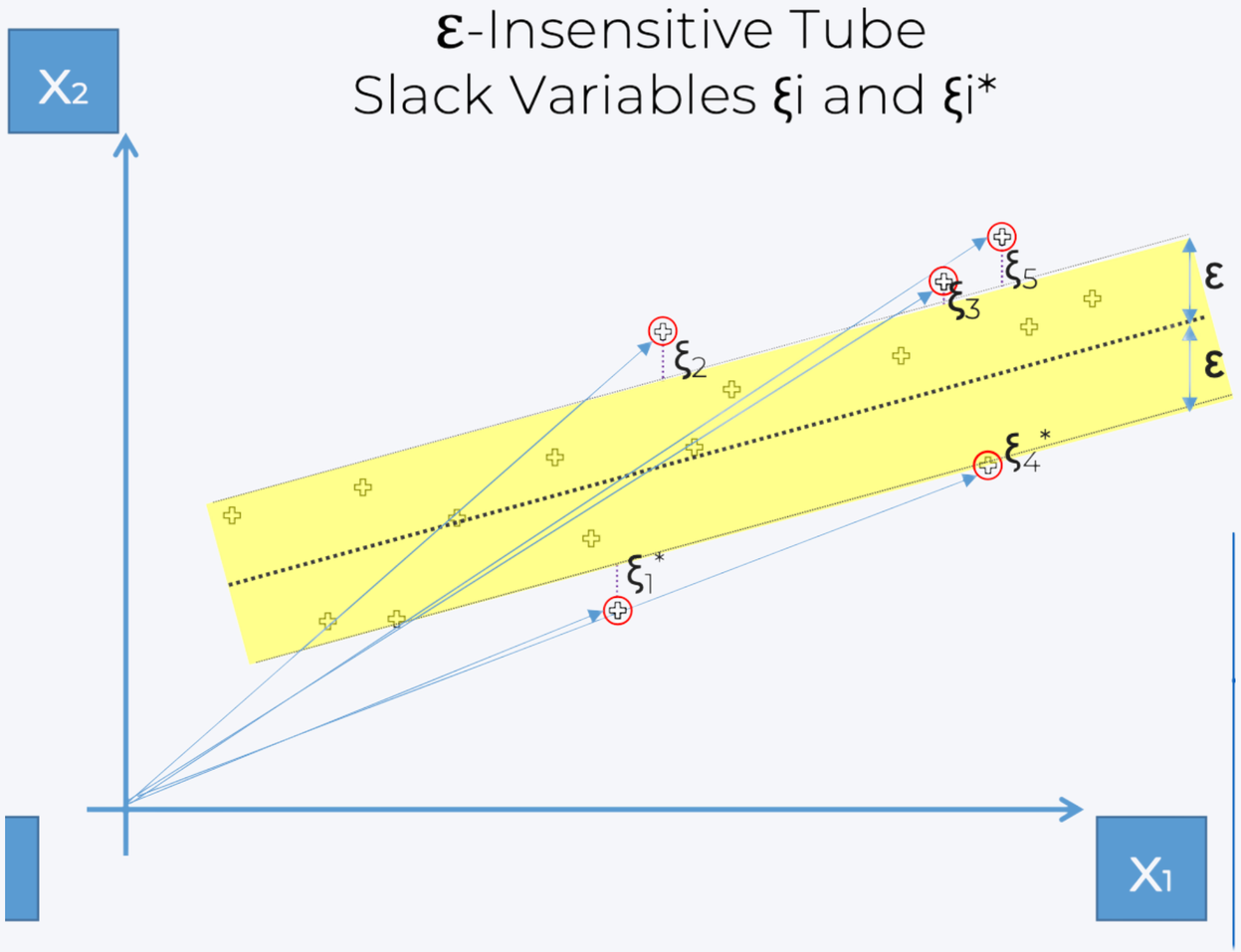


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

What is SVR ?



$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*) \rightarrow min$$

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

Support Vector Regression

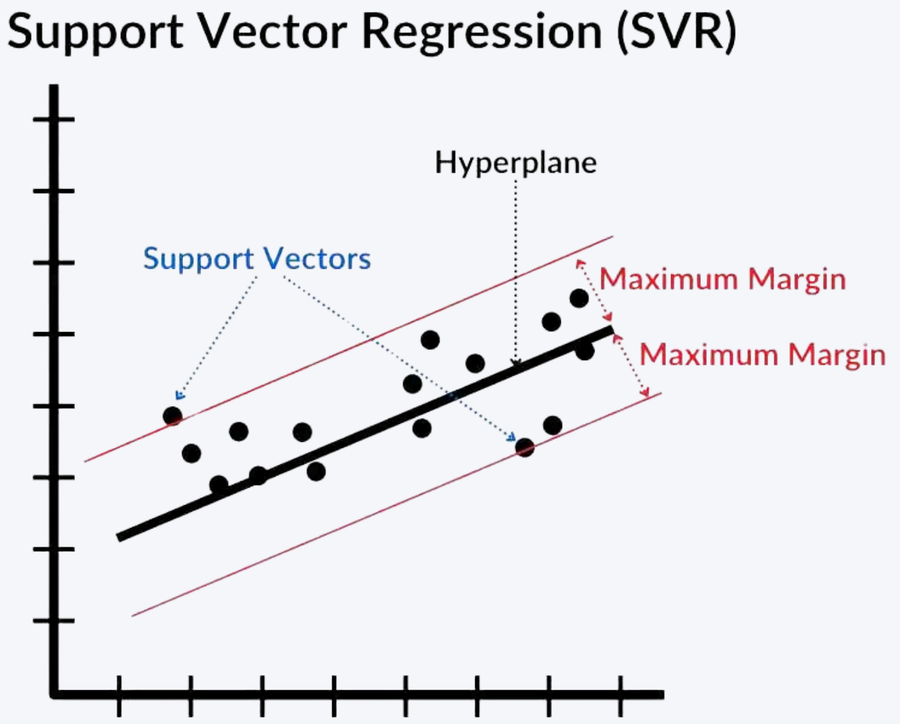
- It finds a **separating hyperplane** with a margin of **±**ε, making it less sensitive to outliers than traditional regression methods.

Support Vectors:

- These are key data points that define the hyperplane's position and help maximize the margin between boundary hyperplanes.

Why Use SVR?

- Handles non-linear data using the **Kernel Trick**
- Uses **quadratic programming** to efficiently find solutions
- Fewer constraints make it computationally efficient



https://spotintelligence.com/2024/05/08/support-vector-regression-svr/

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

Hyperparameters in SVR

Hyperplane:

- **Hyperplanes** are **decision boundaries** that is used to predict the continuous output.
    - The data points on either side of the hyperplane that are closest to the hyperplane are called **Support Vectors.**
    - These are used to plot the required **line** that shows the predicted output of the algorithm.

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

Hyperparameters in SVR

Kernel:

- A **kernel** is a set of mathematical functions that takes **data as input and transform it into the required form.** These are generally used for finding a hyperplane in the **higher** dimensional space.

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

Hyperparameters in SVR

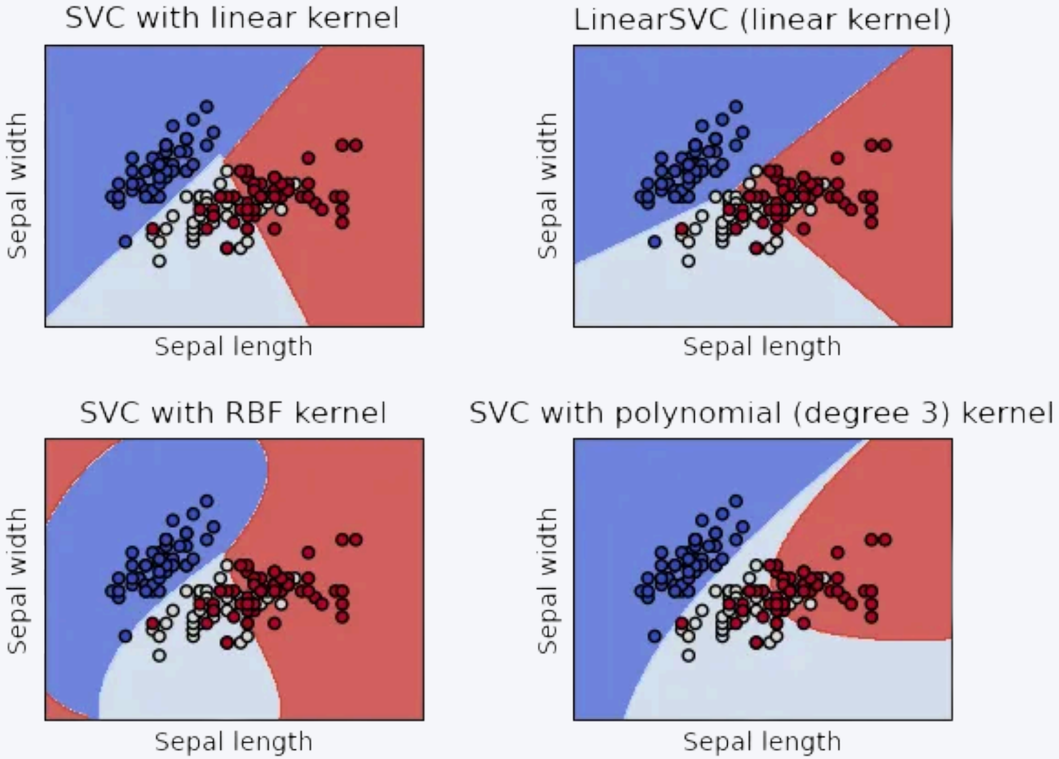
The image compares the decision boundaries of Support Vector Machines (SVM) with different kernels.

SVC with Linear Kernel (Top Left):

- Uses a straight-line decision boundary.
- Suitable for linearly separable data.
- May not perform well with complex data.

LinearSVC (Linear Kernel) (Top Right):

- Another implementation of SVM using a linear kernel.
- Similar decision boundary to SVC with a linear kernel.
- May differ slightly in optimization techniques.



https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

Hyperparameters in SVR

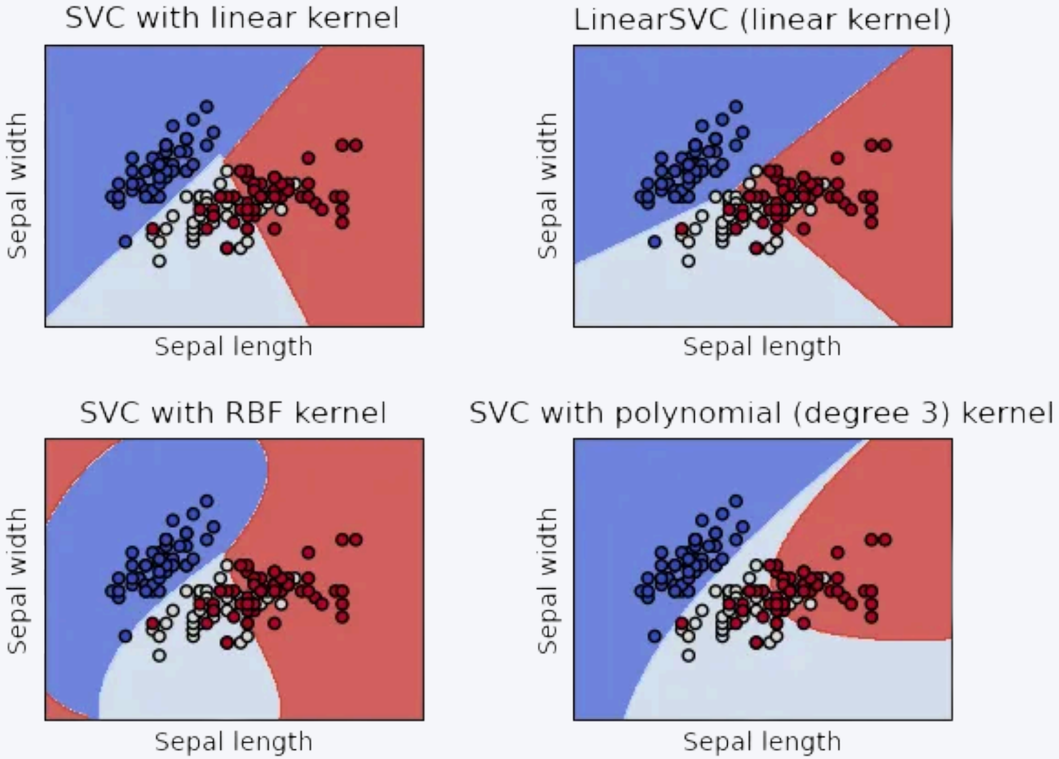
The image compares the decision boundaries of Support Vector Machines (SVM) with different kernels.

SVC with RBF Kernel (Bottom Left):

- Uses a Radial Basis Function (RBF) kernel to create a nonlinear decision boundary.
- More flexible in handling complex patterns.
- Can separate non-linearly distributed data better.

SVC with Polynomial (Degree 3) Kernel (Bottom Right):

- Uses a polynomial function to create a nonlinear decision boundary.
- Can model more complex relationships than a linear kernel.
- May be more sensitive to parameter tuning.



https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

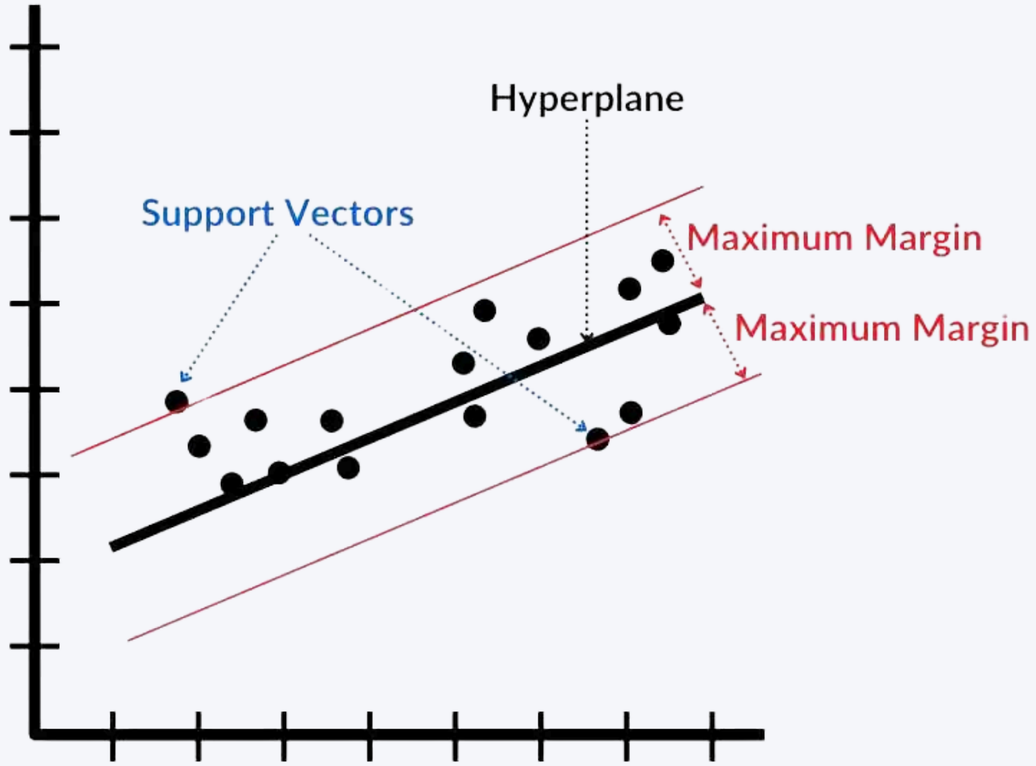
Support Vector Machine

Hyperparameters in SVR

Boundary Lines:

These are the two lines that are drawn around the hyperplane at a distance of ε (epsilon).

It is used to create a margin between the data points.



https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Support Vector Machine

أكاديمية طويق
Tuwaiq Academy

Support Vector Regression (SVR) vs. Traditional Regression:

- Unlike **ordinary regression models** that minimize the error between real and predicted values, **SVR** fits the best line within a threshold margin (epsilon tube).
- Only data points outside the **margin** contribute to model training, making it robust against outliers.
- The computational complexity of **SVR** is more than quadratic in the number of samples, making it inefficient for large datasets.

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

أكاديمية طويق
Tuwaiq Academy

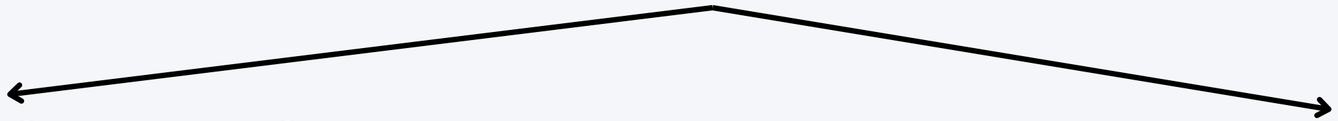
Hands-On Code

SVR Implementation

Decision Tree

Regression

CART



Classification Trees

Regression Trees

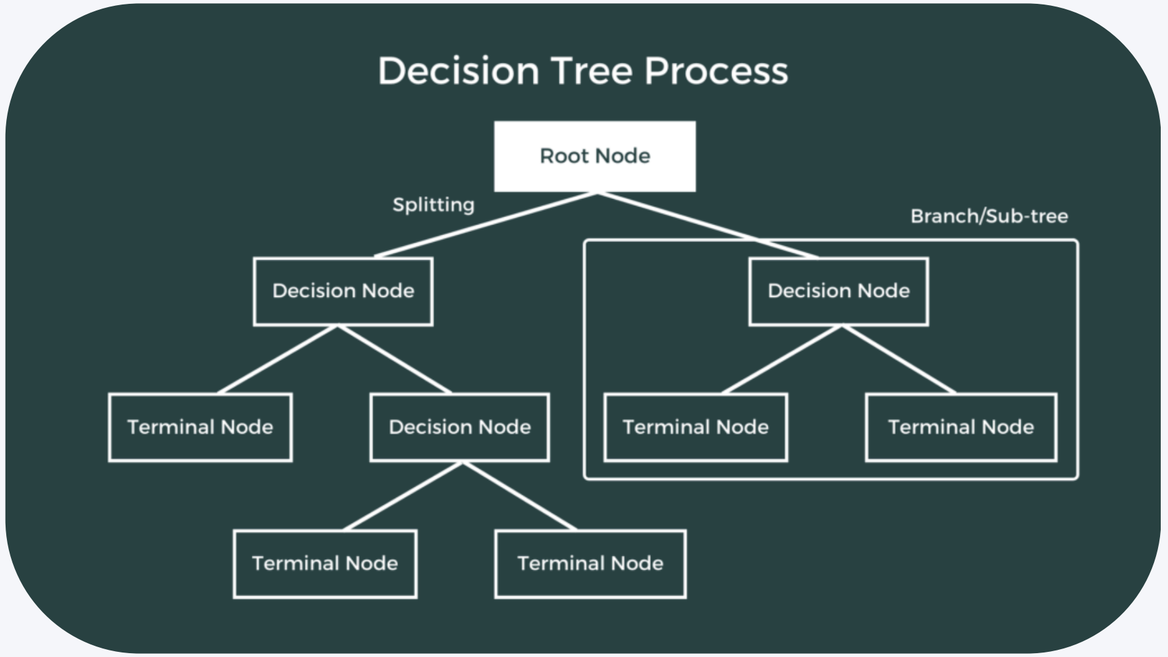
Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

What is Decision Tree Regressors?

- **Decision Tree Regressors** are a powerful, interpretable, and **non-linear** method used widely for regression tasks in machine learning.

- Unlike linear regression, decision trees **partition the feature space in a hierarchical, rule-based way that enables them to capture complex, non-linear relationships**.



https://farshadabdulazeez.medium.com/understanding-decision-tree-regressor-an-in-depth-intuition-a1d3af182efd

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Core Concept of Decision Tree Regressor

A Decision Tree Regressor is a type of **supervised learning algorithm** that uses a flowchart-like structure to predict a **continuous target variable based on decision rules inferred from the data features.** The model continuously splits data into subsets, based on the features that result in the lowest prediction error, forming a tree-like structure where:

- Each **internal** node represents a **decision rule** on a feature.
- Each **branch** represents the **outcome** of a decision.
- Each **leaf node** provides the **predicted value** (usually the average of values in that subset of data).

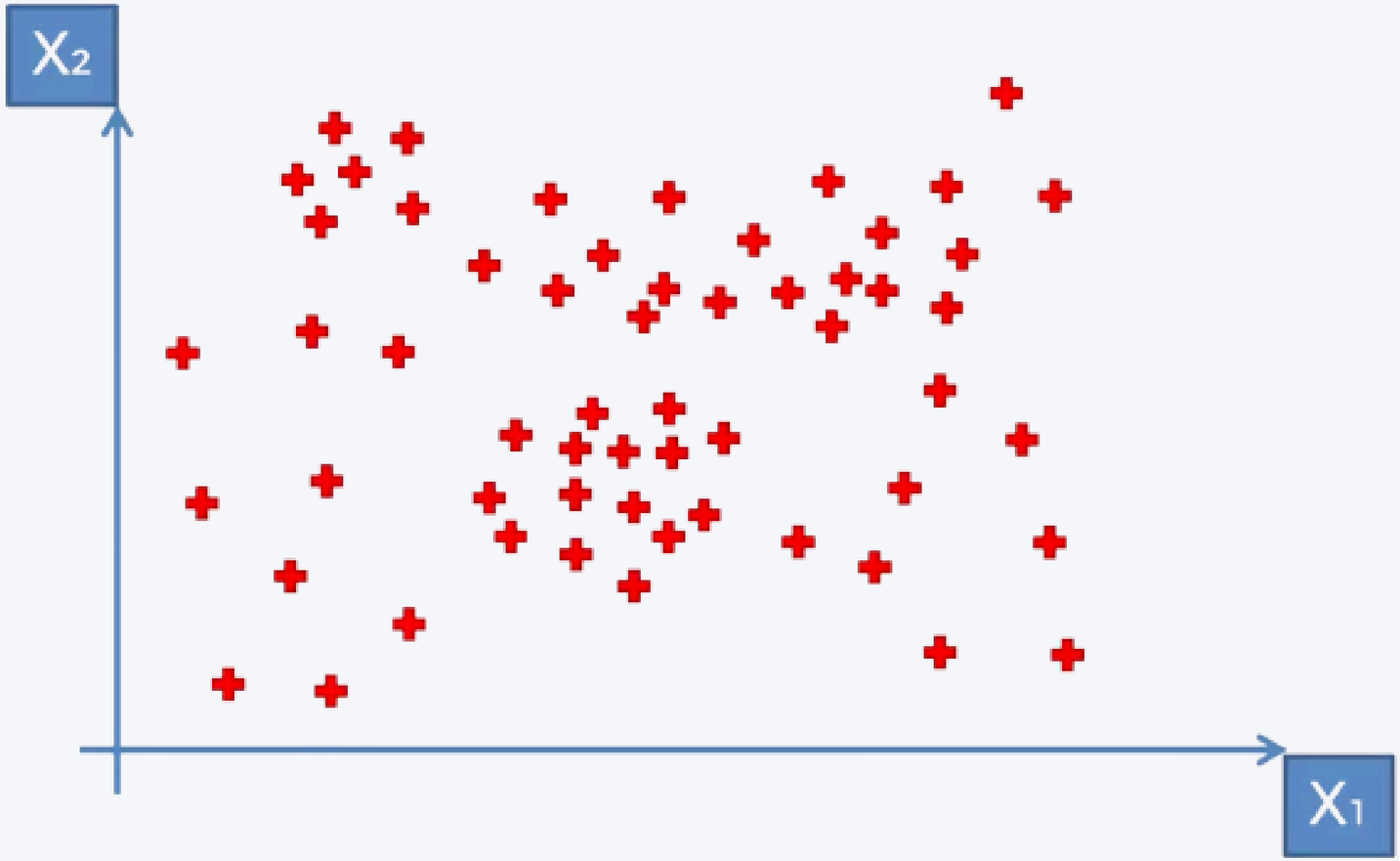
This structure allows the model to capture non-linear relationships in the data by focusing on minimizing errors through the splits.

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

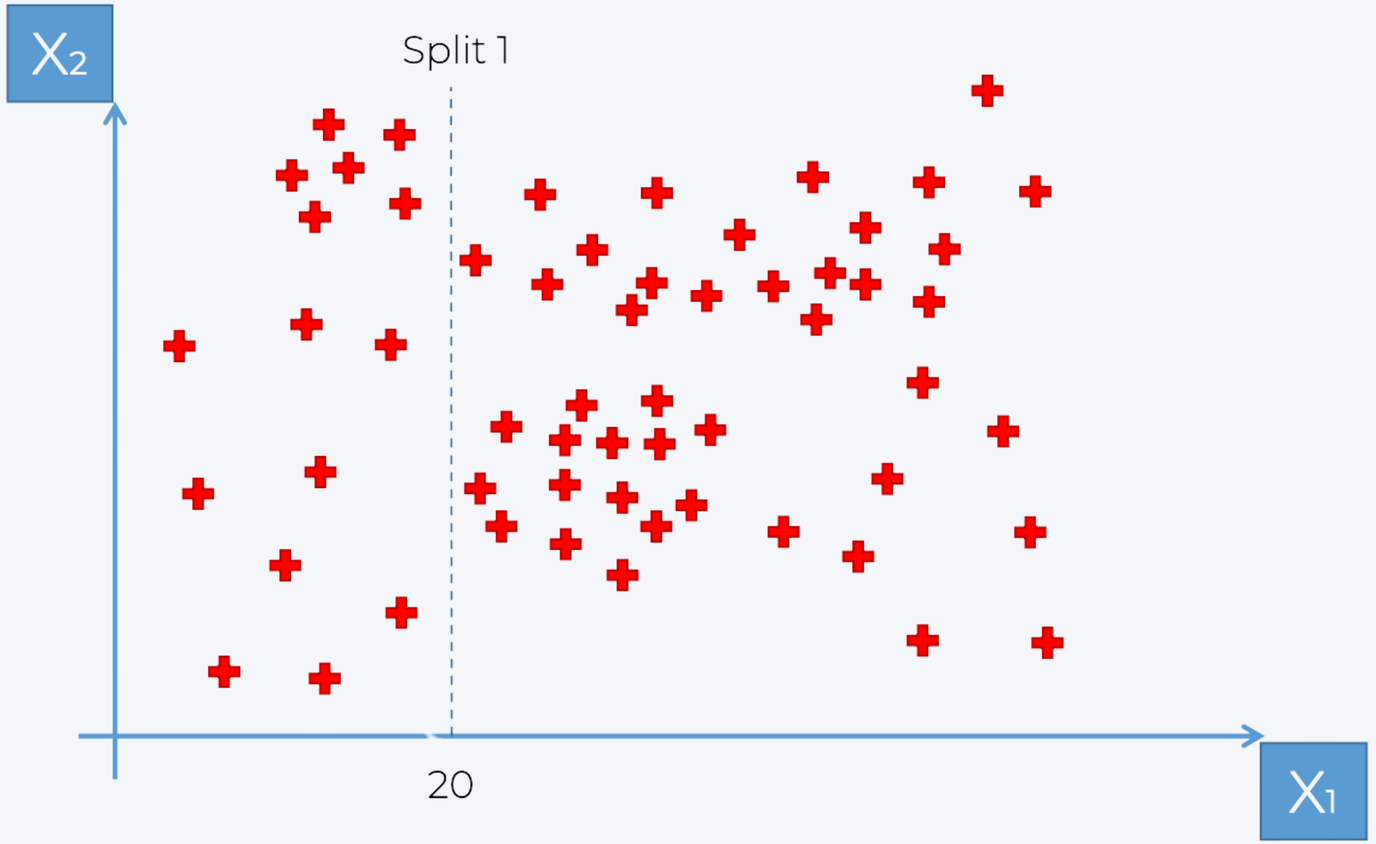


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

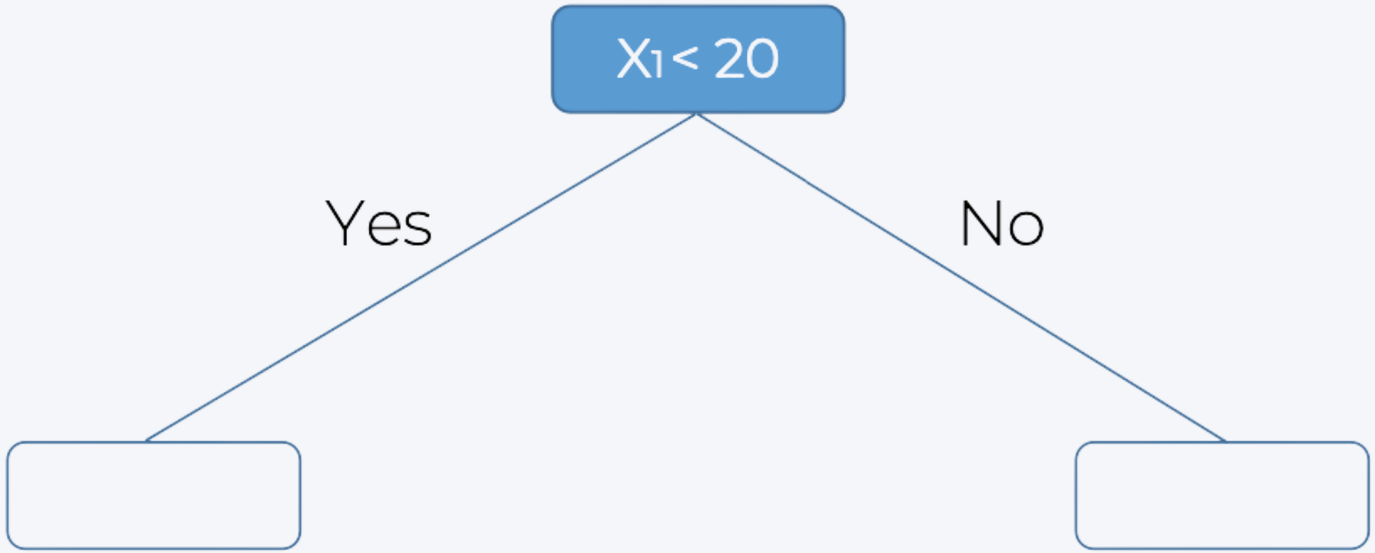


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

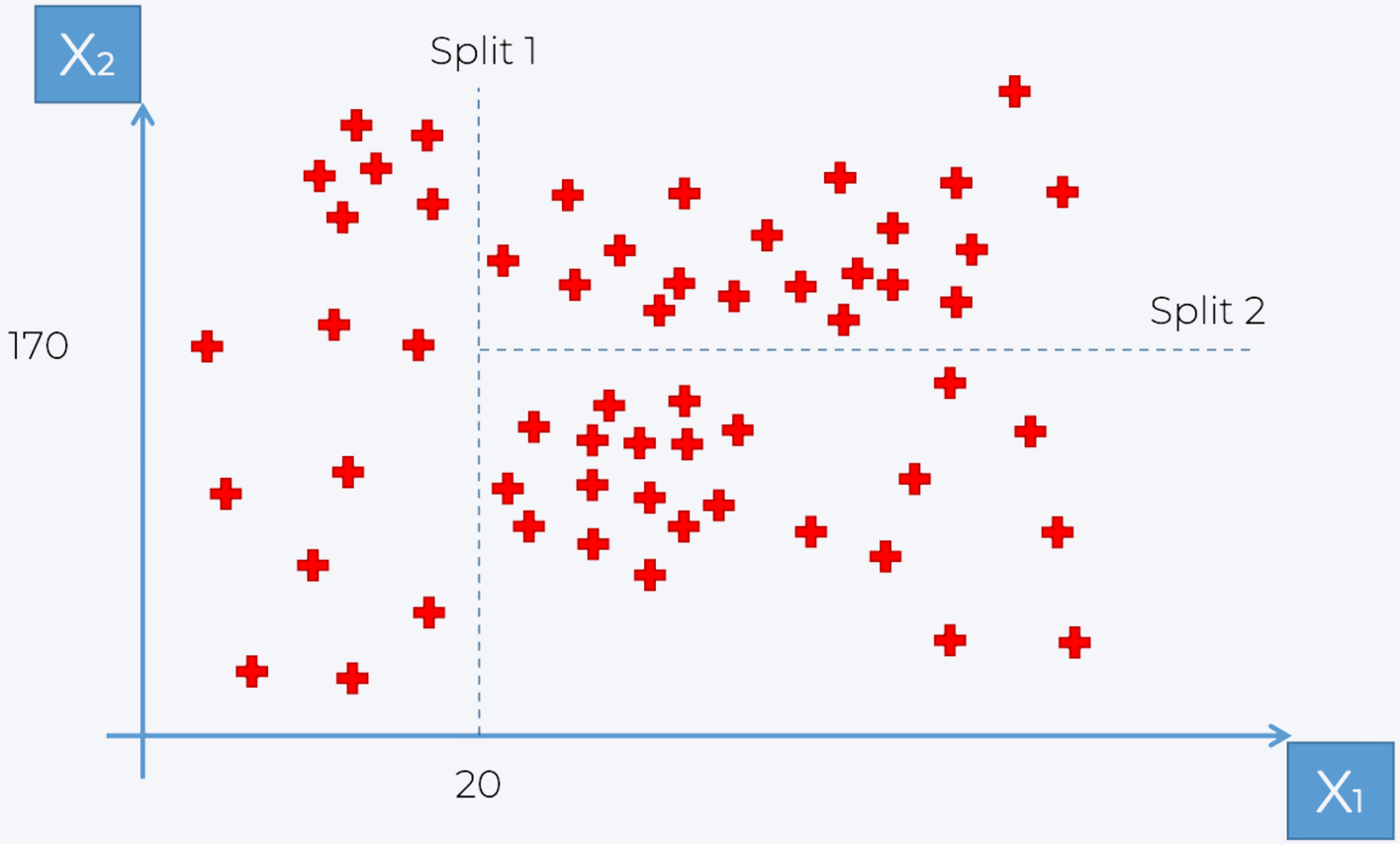


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

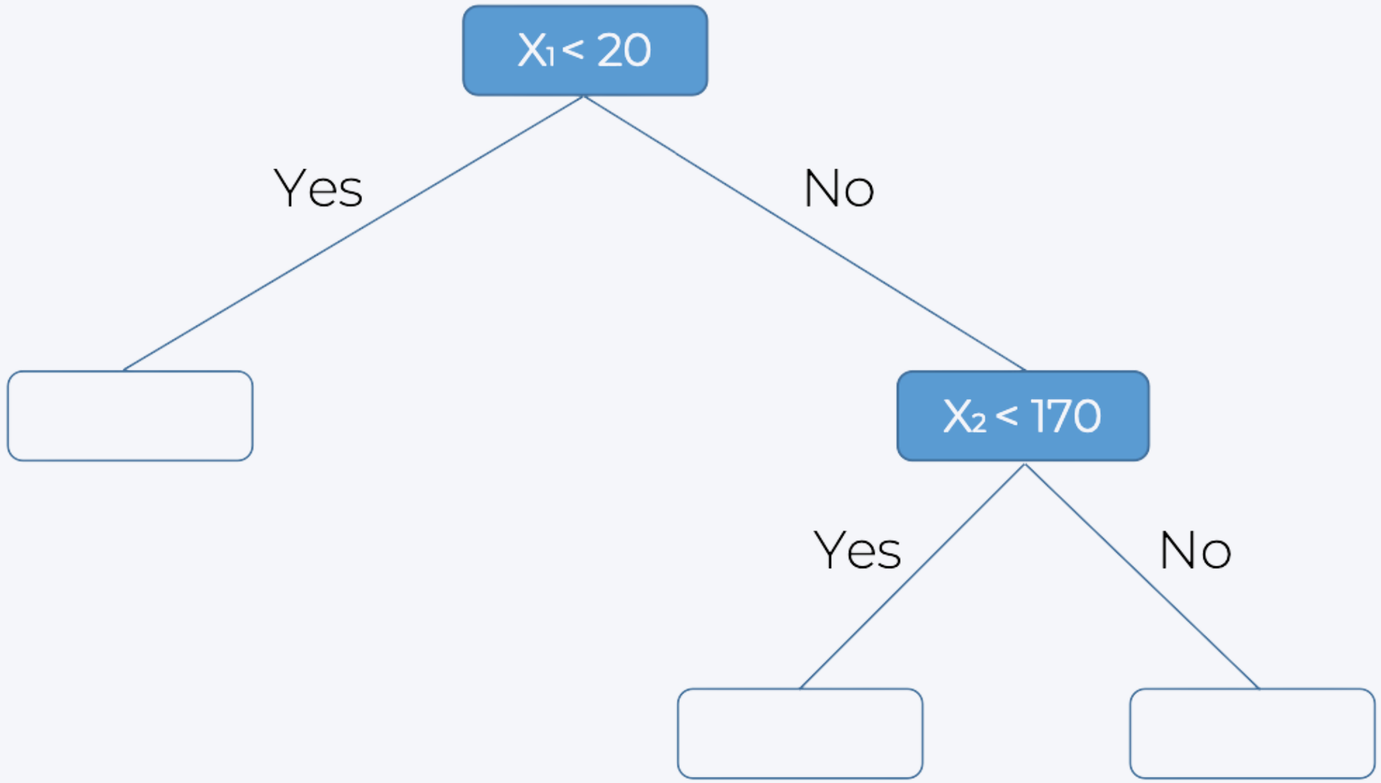


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

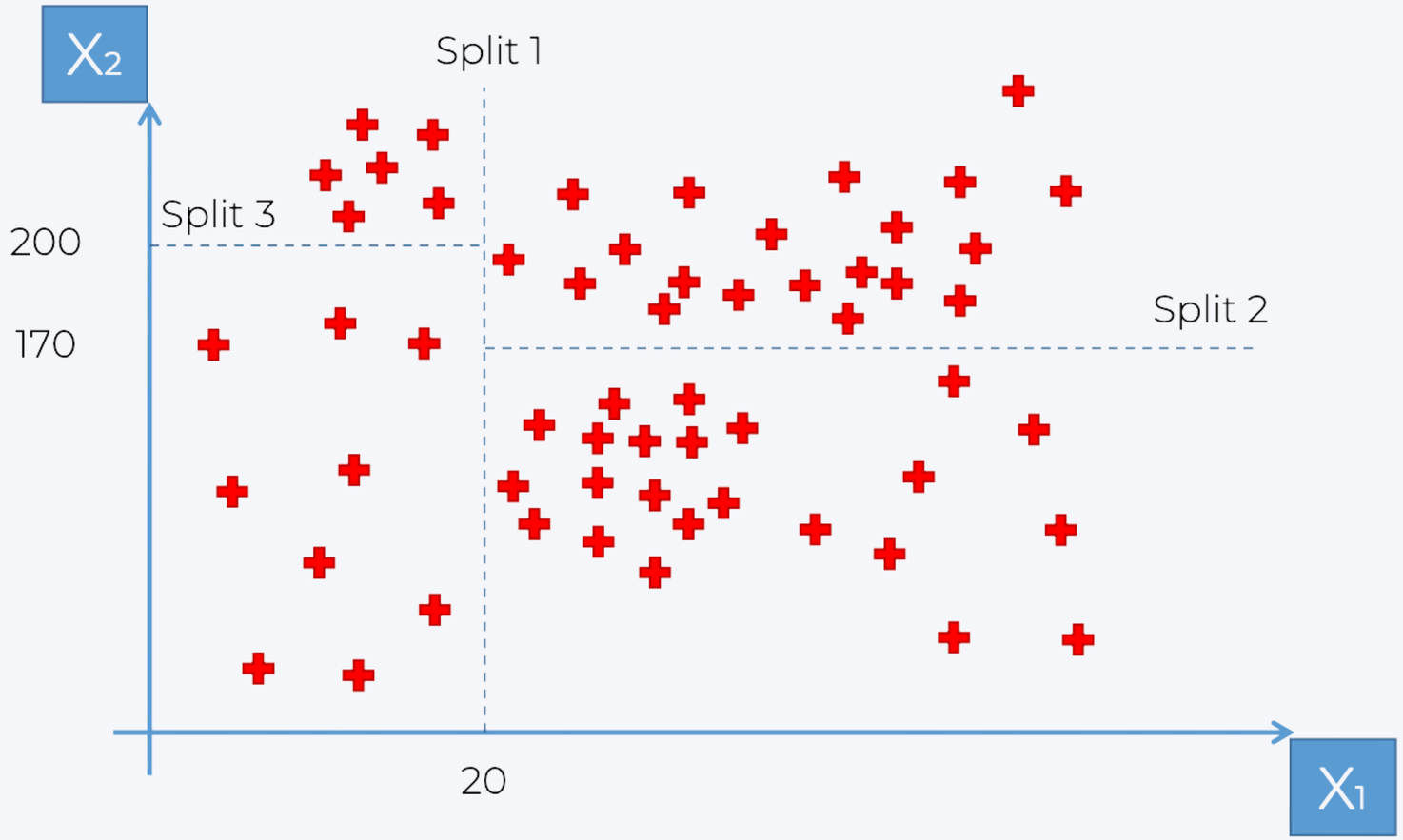


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

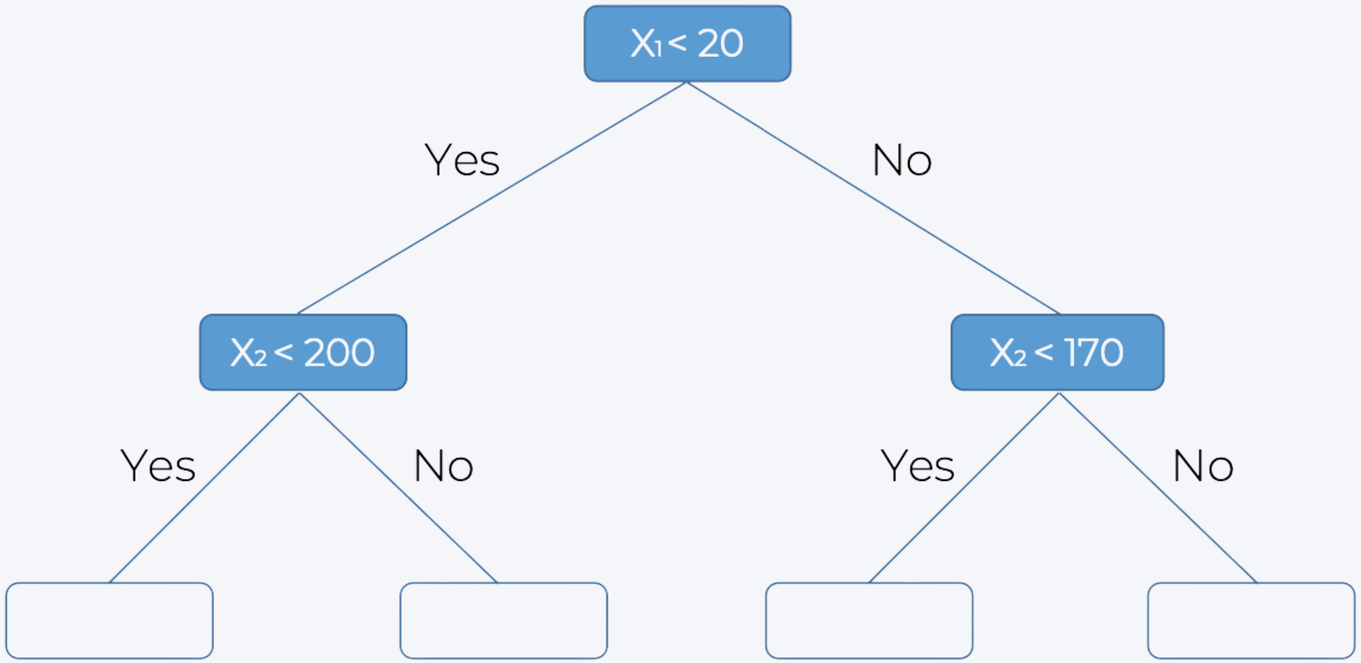


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

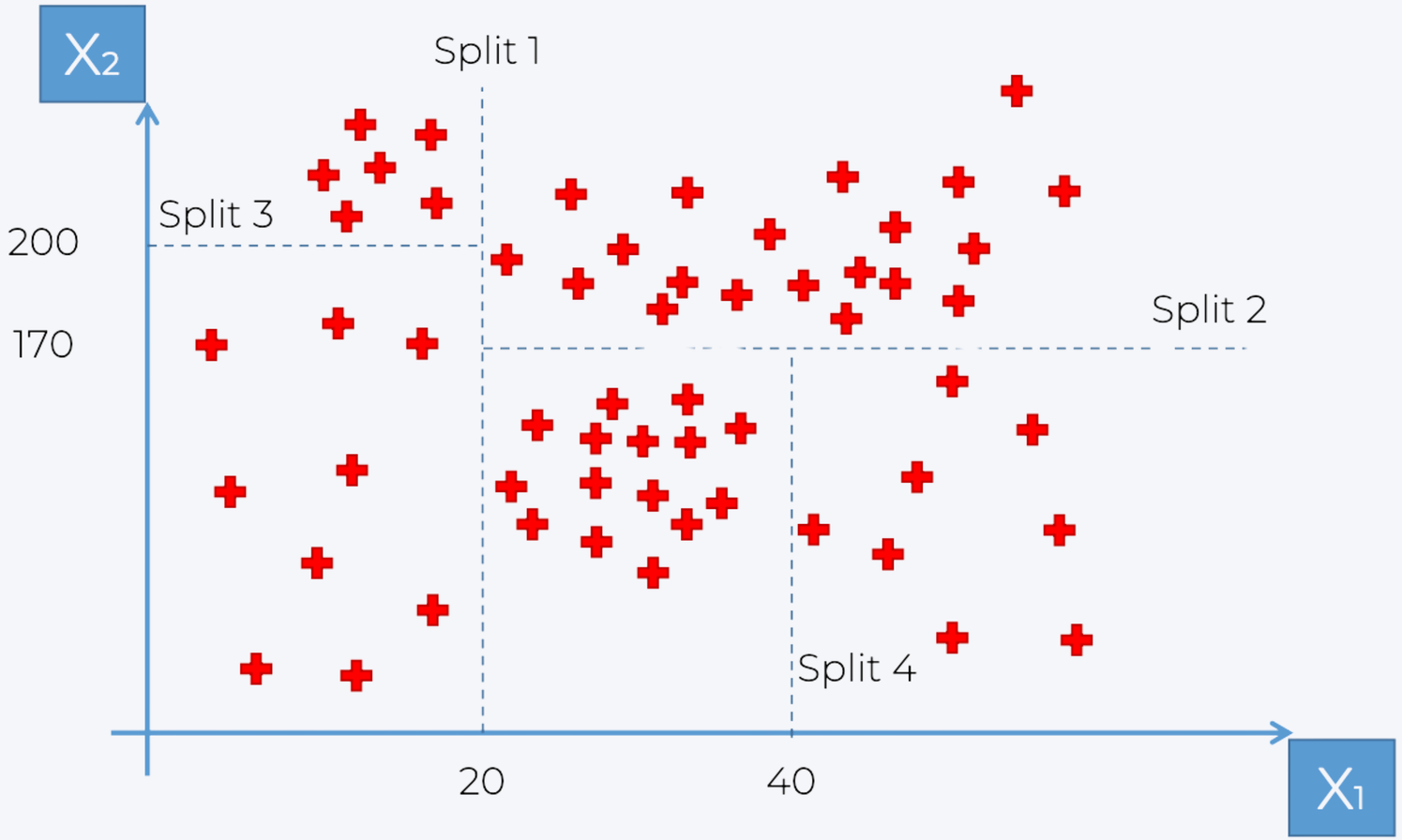


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

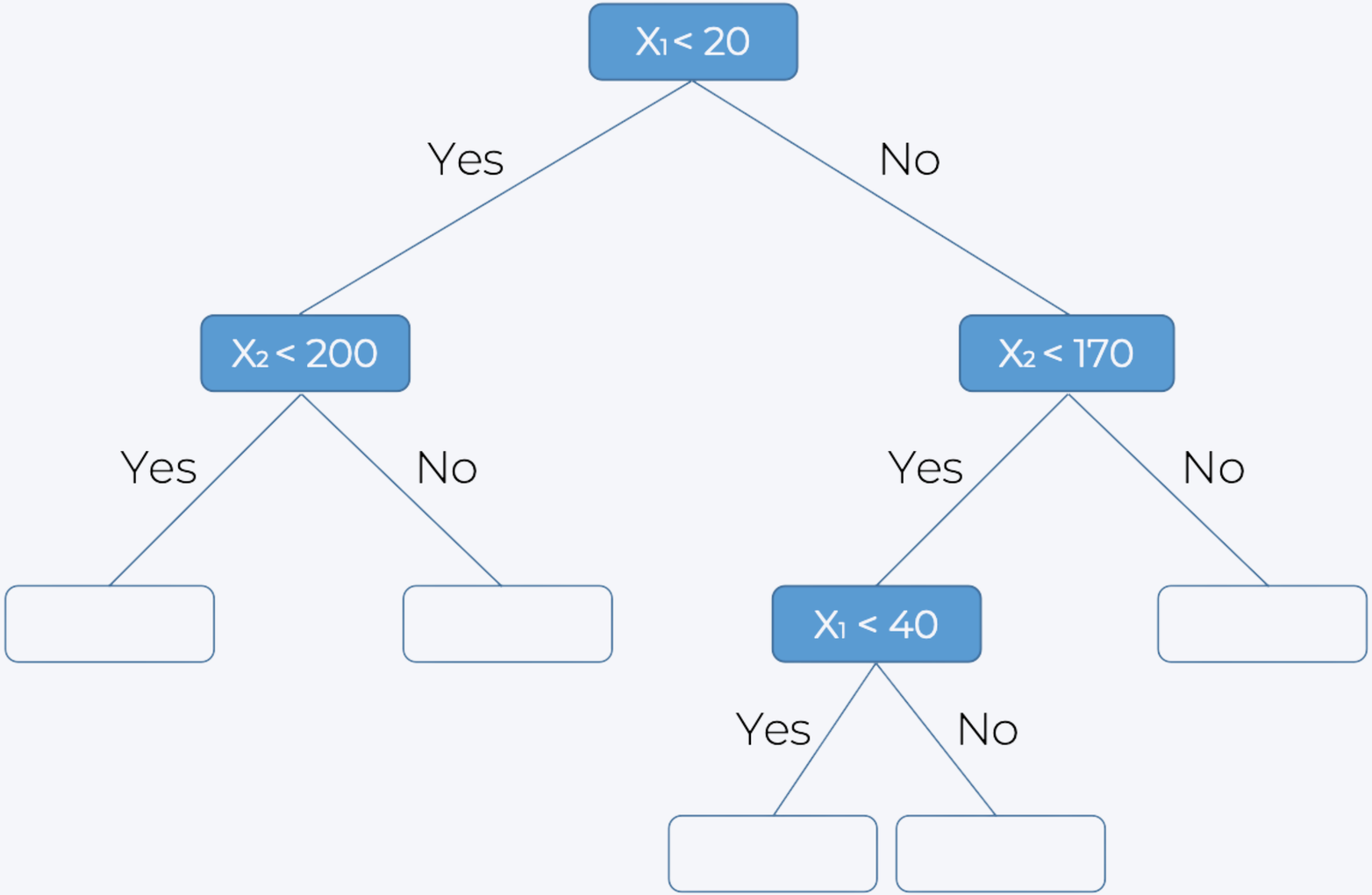


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

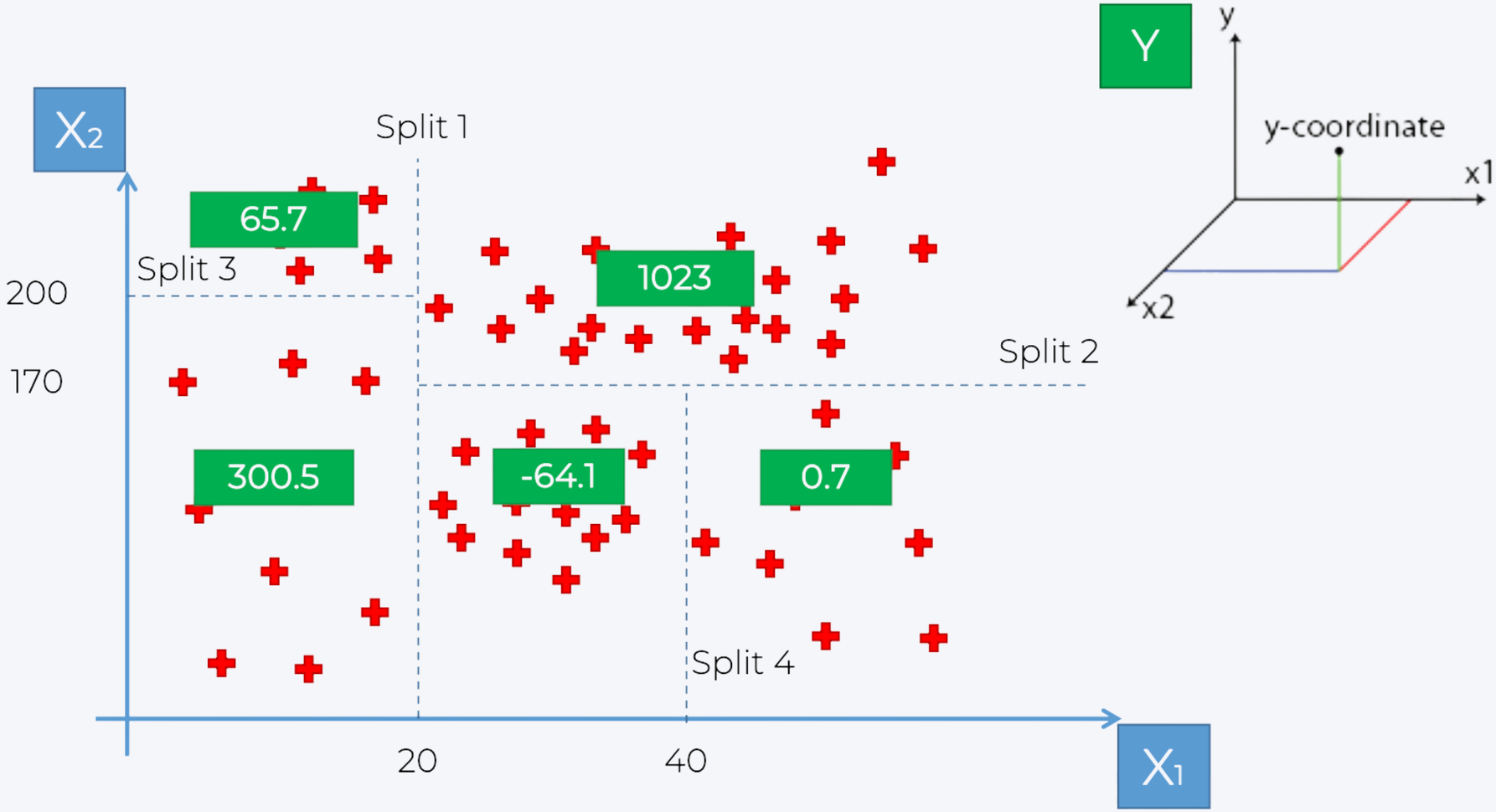


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

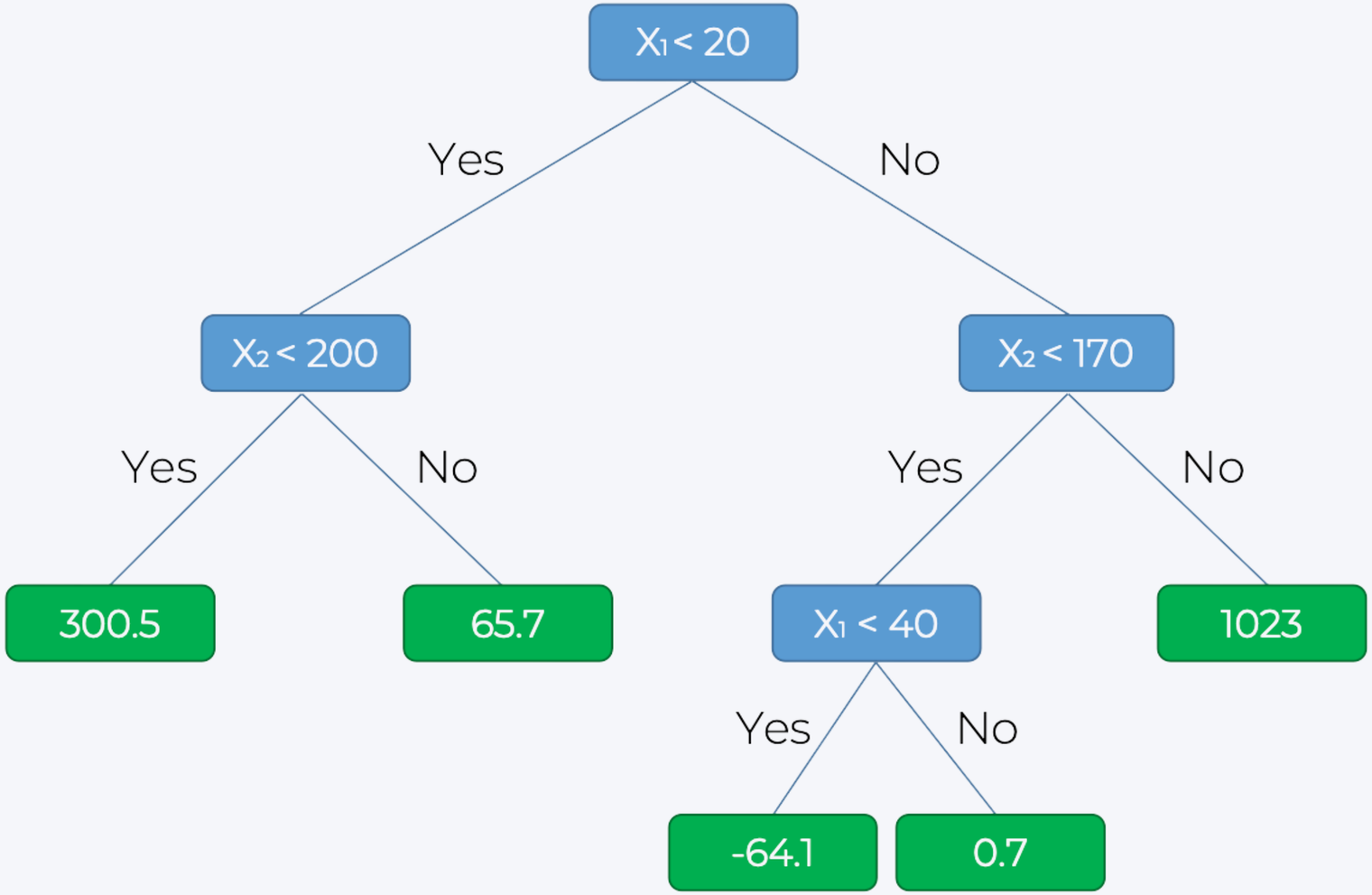


https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm



https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

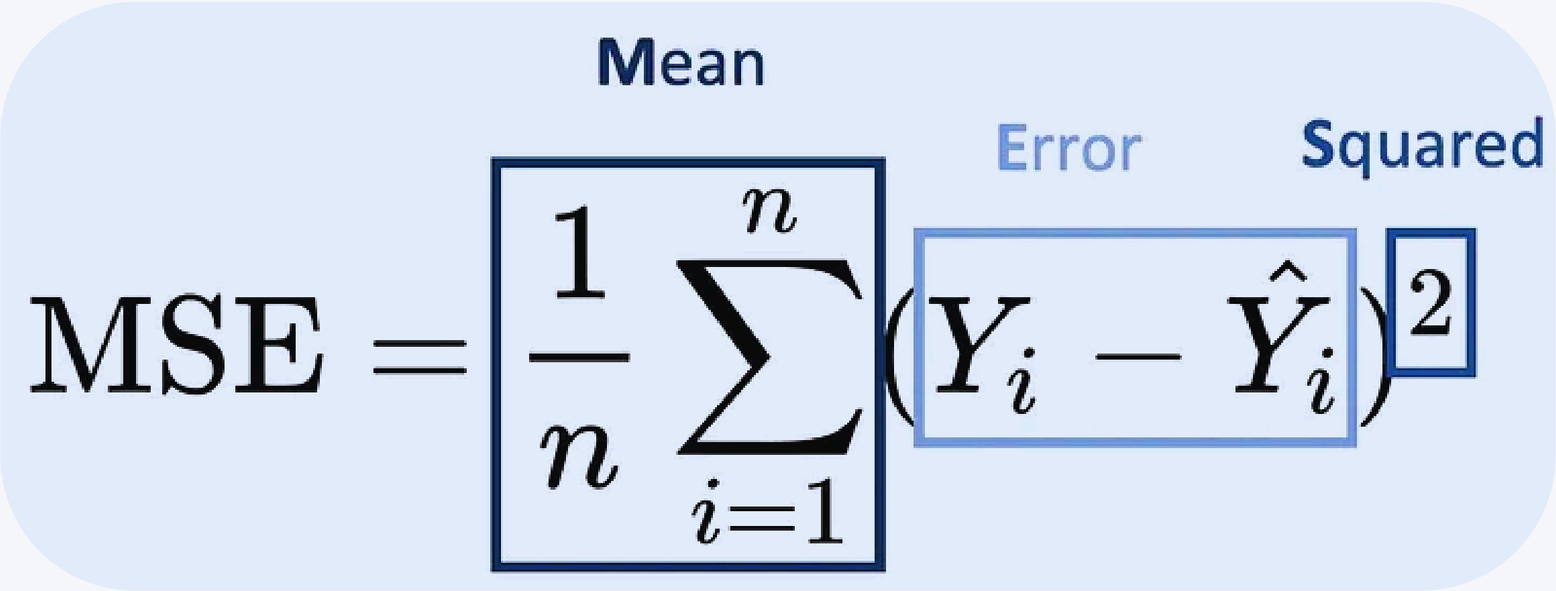
أكاديمية طويق
Tuwaiq Academy

How Splitting Happens:

The most critical step in decision tree regression is determining where to split the data at each node. Here's how it works:

a. Mean Squared Error (MSE) as the Splitting Criterion:

To evaluate potential splits, Decision Tree Regressors typically use **Mean Squared Error (MSE)** as the criterion for **minimizing error**. When considering a split at a **node**, the model aims to reduce the MSE of the target variable, calculated as:



https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Decision Tree Regression

أكاديمية طويق
Tuwaiq Academy

How Splitting Happens:

Start at a node with all data points
The tree considers different possible values to split the data into two parts (left & right).

For each possible split, calculate MSE for both child nodes:

- **MSE_L →** Mean Squared Error for the left child (points below the split).
- **MSE_R →** Mean Squared Error for the right child (points above the split).
- Compute the **weighted average of MSE** for both child nodes:
- Find the split that minimizes MSE_split

The tree chooses the split where MSE_split is the smallest, meaning the data is best divided into two groups.

Mean

Error

Squared

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

$$MSE_{\text{split}} = \left(\frac{N_L}{N} \times MSE_L\right) + \left(\frac{N_R}{N} \times MSE_R\right)$$

NL = Number of points in the left child.
NR = Number of points in the right child.
N = Total number of points at the current node.

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Hands-On Code

Decision Tree Implementation

Random Forest Regression

Random Forest Regression

أكاديمية طويق
Tuwaiq Academy

Algorithm

- STEP 1: Pick at random K data points from the Training set.
- STEP 2: Build the Decision Tree associated to these K data points.
- STEP 3: Choose the number **N** tree of trees you want to build and repeat STEPS 1 & 2
- STEP 4: For a new data point, make each one of your Ntree trees predict the value of Y to for the data point in question, and assign the new data point the average across all of the predicted Y values

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Random Forest Regression

أكاديمية طويق
Tuwaiq Academy

Random Forest Regression Vs. Decision Tree

Feature	Decision Tree 🌳	Random Forest 🌲🌲🌲
Definition	A single tree that splits data using MSE (for regression).	An ensemble of multiple Decision Trees that work together.
How it Works	Finds the best splits in data to minimize MSE.	Picks random subsets of data and builds multiple trees, then averages their predictions.
Overfitting Risk	High risk ⚠ (memorizes data).	Low risk ✅ (averages multiple models for better generalization).
Accuracy & Stability	Less stable, small changes in data can change the tree.	More stable & accurate, reduces variance by averaging predictions.
Final Prediction	The output of one tree.	The average of multiple tree outputs (for regression).

https://www.researchgate.net/publication/344389401_DS-Regression-04-SVR

Hands-On Code

Random Forest Implemenration