Support Vector Regression

Definition - Support Vector Regression (SVR) is a type of regression algorithm that is an extension of Support Vector Machines (SVM). It is designed to predict continuous values while maintaining the key principles of SVM by **finding the best boundary (or hyperplane) that fits within a margin of tolerance**. In essence, SVR aims to fit a regression model by minimizing the prediction error within a specified threshold (ε).

Feature Scaling - **Feature scaling is essential** in Support Vector Regression (SVR) because the model relies on distance-based calculations to determine the best-fit hyperplane. Without scaling, features with larger ranges dominate these distance measurements, leading to biased results. Additionally, SVR often uses kernel functions (like RBF or polynomial), which are sensitive to the relative distances between data points, so unscaled features can negatively impact the model's ability to capture patterns. Lastly, the regularization term in SVR, which controls model complexity, may behave inconsistently if features are not scaled, resulting in suboptimal generalization.

How It Works

1 - **Hyperplane and Margin**: SVR constructs a hyperplane in a high-dimensional feature space. SVR attempts to fit the hyperplane within a margin of tolerance ε around the actual data points. This margin defines how much error is tolerable for the model

2 - **Error Tolerance (ε-insensitive loss)**: The model ignores errors that fall within the tolerance zone (ε). Only data points that fall outside this zone contribute to the model’s loss function. This makes SVR robust to small errors, allowing it to generalize better to unseen data.

3 - **Support Vectors**: The data points that lie outside the margin of tolerance are called support vectors. These support vectors are critical in defining the position of the hyperplane, as they are the points the model uses to learn the regression function.

4 - **Kernel Trick**: SVR can handle non-linear data by using kernel functions (such as polynomial, radial basis function, etc.). These kernels project the data into higher-dimensional spaces, allowing SVR to learn complex relationships in the data.