44. Support Vector Machines (SVM) - Classification

-- SVM is one of the most popular **Supervised Learning** algorithm, which is used fro classification as well as Regression problems

- With the help of SVM, you can handle both linear and non-linear data
- Its working is similar to logistic regression algo
- **Algo of SVM**:

1. It finds two points (support vectors) in the data (shown in red color in below figure)
2. It passes marginal plan/line (hyperplane) from these support vectors (dotted red lines)
3. It measures the distance b/w these two lines
4. Take avearge of the distance (divided by 2)
5. This average is denoted by a separable line (Maximum margin) (solid red line)
6. Prediction is done through this line and decided the new data would go to which category (Splitting of data)
7. The distance (d) b/w 2 vectors should be maximum

No description has been provided for this image

- **Hard Margin:** The algorithm aims to find a hyperplane that perfectly separates the data into two classes w/o any misclassifications.
- **Soft Margin:** The algorithm allows for some misclassifications to find a hyperplane that generalizes better to unseen data and is more robust to outliers

No description has been provided for this image

No description has been provided for this image

In [ ]:

```

```

Types of SVM: There are two different types of SVMs, each used for different things:

1. **Simple SVM:** Typically used for linear regression and classifications problems.
2. **Kernel SVM:** Has more flexibility for non-linear data b/c you can add more features to fit a hyperplane instead of a two-dimensional space.

- Kernel SVM is used when our data is not linearly separable.
- It modifies our data

Kernel Functions:

- Kernel functions play a crucial role in transforming input into a higher-dimensional space.
- The primary purpose of kernel functions is to allow SVMs to handle non-linearly separable data by implicitly mapping the input data into a higher-dimensional feature space where linear separation may be more feasible.
- This transformation is done w/o explicitly calculating the coordinate points in that higher-dimensional space.

Kernel Functions in SVM

1. **Linear Kernel**:

$$K(x_i, x_j) = x_i^T x_j$$

2. **Polynomial Kernel**:

$$K(x_i, x_j) = (\gamma \cdot x_i^T x_j + r)^d$$

3. **Gaussian Radial Basis Function (RBF) Kernel**:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right)$$

4. **Sigmoid Kernel**:

$$K(x_i, x_j) = \tanh(\gamma \cdot x_i^T x_j + r)$$

Description of Symbols

- (x_i, x_j): Input feature vectors.
- (x_i^T x_j): Dot product between vectors (x_i) and (x_j).
- (\gamma): Scaling factor (often related to (\sigma) in the Gaussian RBF kernel as (\gamma = \frac{1}{2\sigma^2})).
- (r): Constant term (bias).
- (d): Degree of the polynomial in the Polynomial Kernel.
- (|x_i - x_j|): Euclidean distance between (x_i) and (x_j).
- (\exp): Exponential function.
- (\tanh): Hyperbolic tangent function.

In [ ]:

```

```