PAGE NO :
DATE :

Support Vector Machine (SVM)

→ Input:
Dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$
Test dataset X_{test}, Regularization parameter C, Max. no of iter. max_iter

→ Output:
Predicted class labels y pred
Classification accuracy.

1. Data Loading and Preprocessing
- Load the Iris dataset.
- Apply Z-score normalization to standardize features $x' = \frac{x - \mu}{\sigma}$

2. Split the Dataset.
- split the data into
Training set (70)%, Test set (30%)

3. Initialize The SVM classifier
- set the following parameters
C: Regularization constant
$K(x, x') = x \cdot x'$

4. One vs Rest Training strategy

For each class c in the set of unique classes:

Convert labels into binary format.

PAGE NO :
DATE :

$$y_{binary} = \begin{cases} 1 & \text{if } y = c \\ -1 & \text{otherwise} \end{cases}$$

Train a binary SVM classifier using the simplified SMO algorithm.

5. Binary SVM Training

Initialize :
$\alpha = \vec{0}$: Lagrange multipliers
$b = 0$: Bias term.

Repeat for max iter iterations :
For each training sample i :
(i) Randomly select another index $j \neq i$
(ii) Compute prediction errors E_i & E_j
(iii) Save old values α_i, α_j
(iv) Compute bounds L, H

if L = H, continue

(v) Compute

$$\eta = 2K(x_i, x_j) - K(x_i, x_i) - K(x_j, x_j)$$

If $\eta \geq 0$, skip update.

vi) Update α_j :

$$\alpha_j = \alpha_j + \frac{y_j(E_i - E_j)}{\eta}$$

vii) Clip $\alpha_j \in [L, H]$.
viii) Update α_j

(ix) Compute :

PAGE NO :

DATE :

$b_1 = b - E_i - y_i(\alpha_i - \alpha_i^{old}) K(x_i, x_i) - y_j(\alpha_j - \alpha_j^{old}) K(x_i, x_j)$, compute b_2

4) Update the bias term b

Check for convergence. $\|\alpha - \alpha_{prev}\| < 10^{-5}$

6. Prediction phase.

- For each test sample x:
- For each trained binary classifier

Compute decision score:

$$f(x) = \sum_j a_j y_j K(x, x_j) + b$$

Store the score, predict the class with the maximum decision score.

7. Evaluation

- Compare predicted labels y_{pred} with true labels y_{test}.

- Calculate accuracy.

$$\text{Accuracy} = \frac{\text{No of correct predictions}}{\text{Total test samples}} \times 100$$