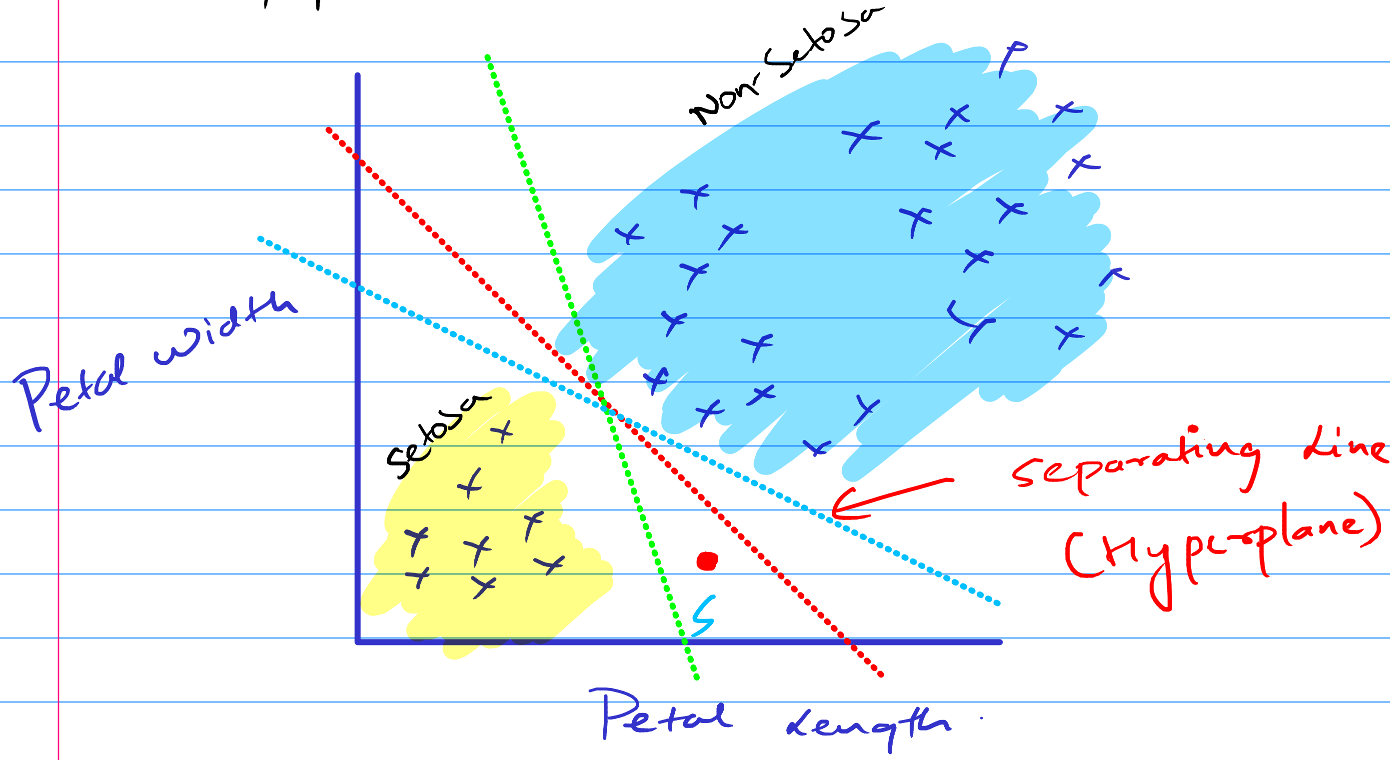
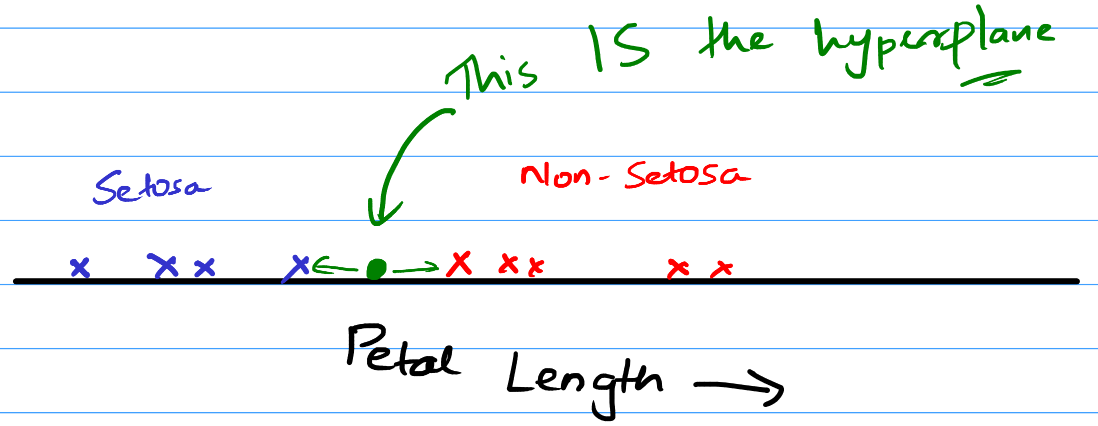
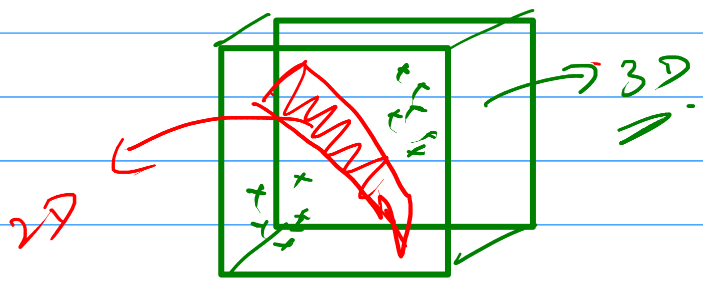
Support Vector Machines.

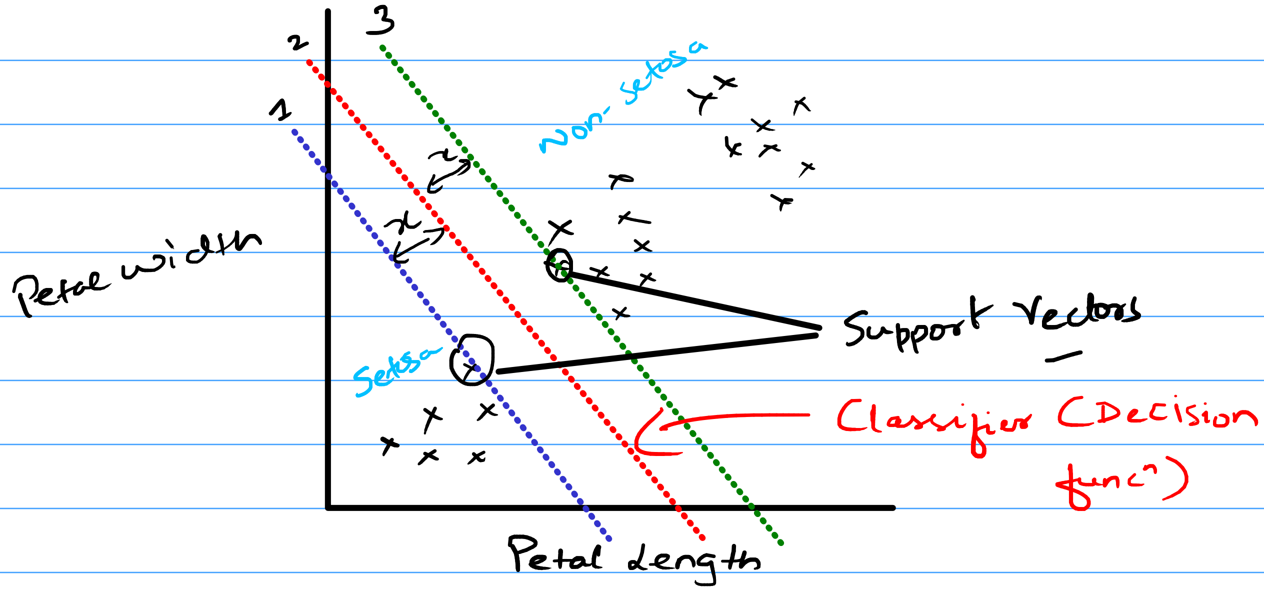


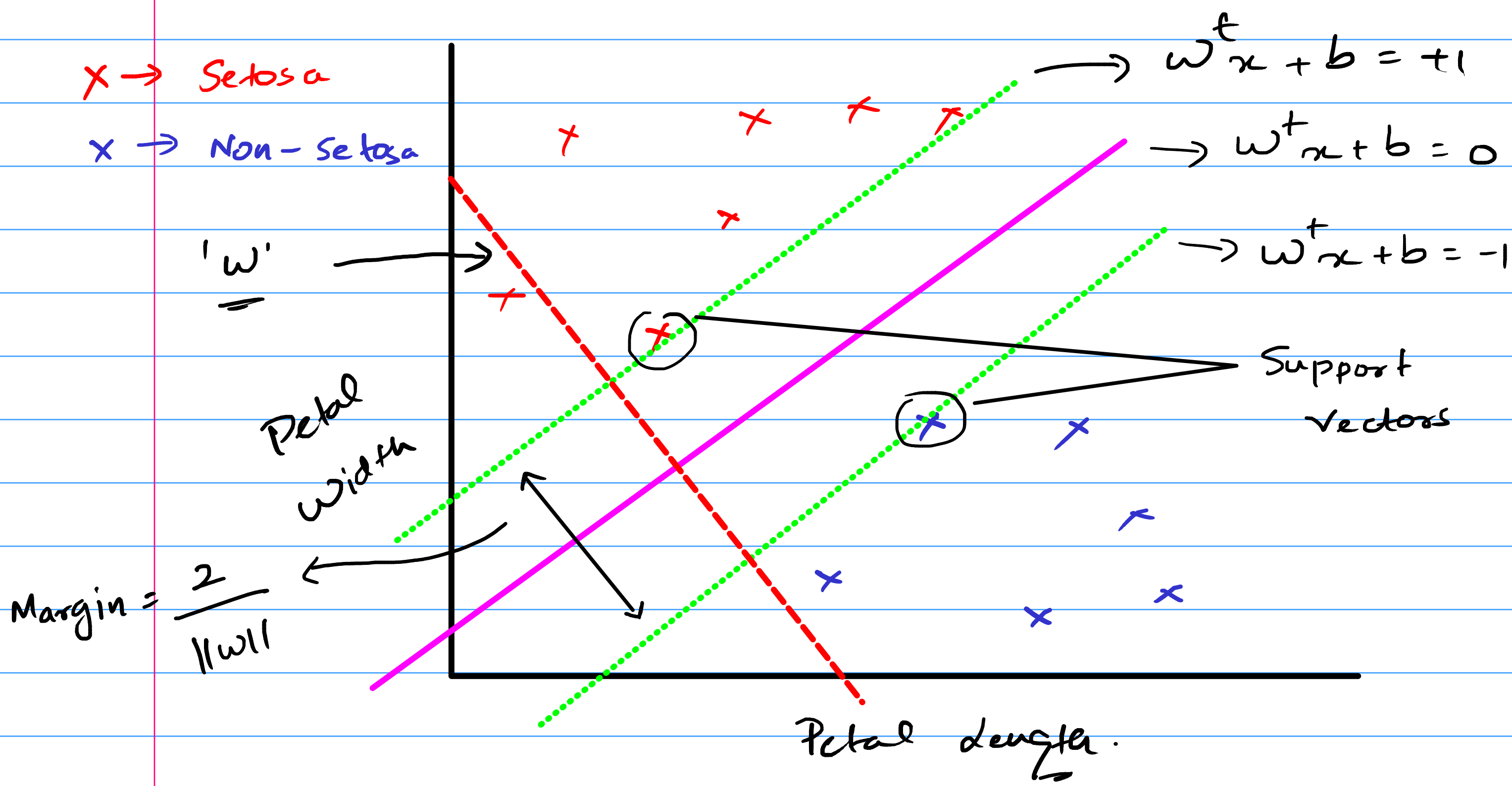
The goal of SVM can be defined as to find an optimal hyperplane that separates the data into classes.

Hyperplane: **is a subspace of one dimension less than its ambient space.**









Decision Boundary Expression:

$$f(x) = \sum_{i=1}^{n} w_i^T x_i + b$$

where $i = 1 - n$ data points

x = training or test patterns

w = is a weight vector Normal to the hyperplane (red dashed line)

b = bias

The result of this func^n → +ve / → −ve based on the predicted class of the test data point.

Expression of the Hyperplanes:

Decision hyperplanes divides the two classes of data → +ve / → −ve

All the data points that fall on the hyperplane has a value 0, when substituted by the expression " $w_i^T x_i + b$ ".

if you move on the +ve side ⇒ its value ↑ from 0.
if you move on the −ve side ⇒ its value ↓ from 0.

Understanding the Weight Vector "w".

$$w = \sum_{i=1}^{n} a_i y_i x_i ,$$

$$y_i = (-1, +1)$$

where a_i value is close to 0 for most points except support vectors.

Let "u" represent the new test data point.

Then $w = \sum a_i y_i u$

$$y_i = (-1, +1)$$

"y_i" is -1 for data points that fall on the $-ve$ side of the hyperplane

"y_i" is $+1$ " " " " " " " $+ve$ side " " ".

$\therefore$, $y_i (w_i^T x_i + b)$ is always positive.

Margin Width:

$$= \frac{2}{\|w\|}$$

where $\|w\|$ = length of the vector

$= \sqrt{w^T w}$

Maximum Margin:

is achieved by finding "w" & "b":

$\frac{2}{\|w\|}$ is maximized $\forall \{(x_i, y_i)\}$,

$y_i (w_i^T x_i + b) \geq 1$

In case of 2 features, the discriminant func^n is

$$f(x) = w_1 x_1 + w_2 x_2 + b$$

When evaluating this equation, the resulting sign divides the 2D space into 2 regions. After training, the SVM provides us with the estimates for w1, w2, and b.

The decision function can thus be constructed as the line, which crosses the discriminant function at the point where f(x) = 0.

The dotted (green) lines can be constructed as the lines crossing the discriminant function at the points where it evaluates to "-1" and "+1" respectively.