VIT BHOPAL
www.vitbhopal.ac.in

# Supervised Learning

## (Introduction to Support Vector Machine)

**Dr. Virendra Singh Kushwah**

**Assistant Professor Grade-II**

**School of Computing Science and Engineering**

**Virendra.Kushwah@vitbhopal.ac.in**

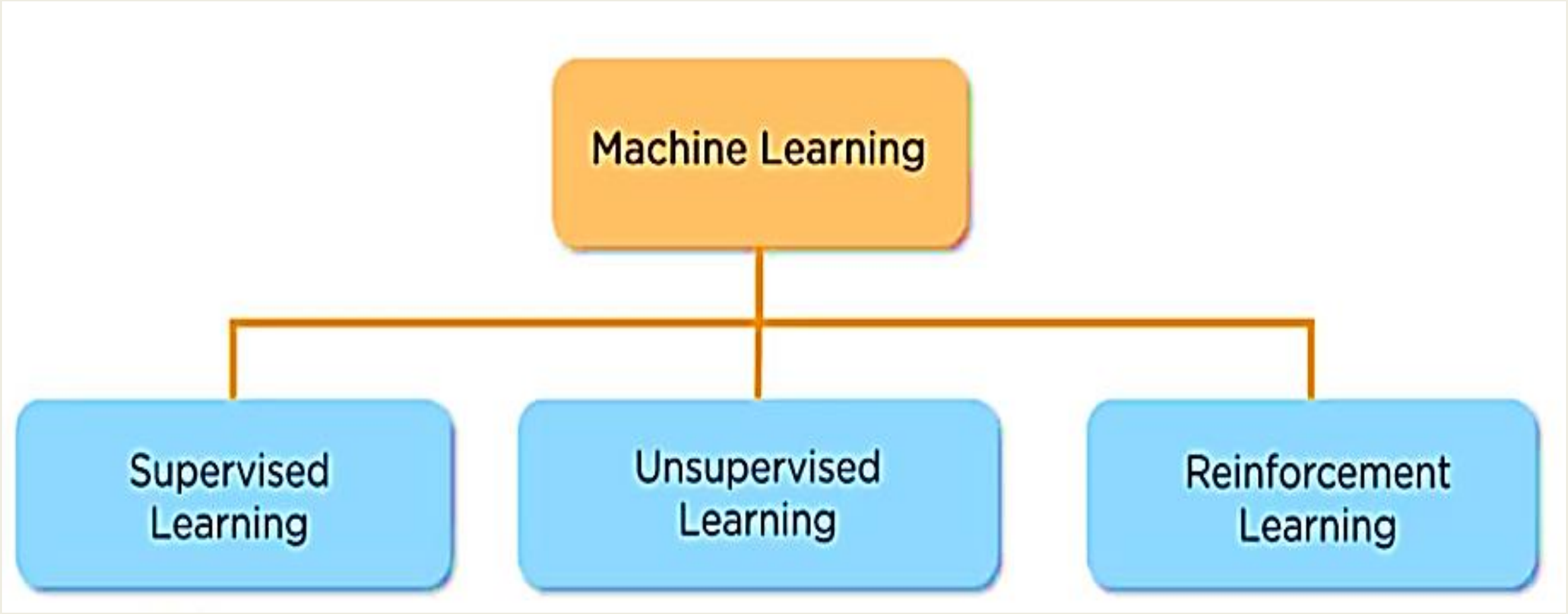
**7415869616**


VIT®
BHOPAL
www.vitbhopal.ac.in

# Points to be discussed

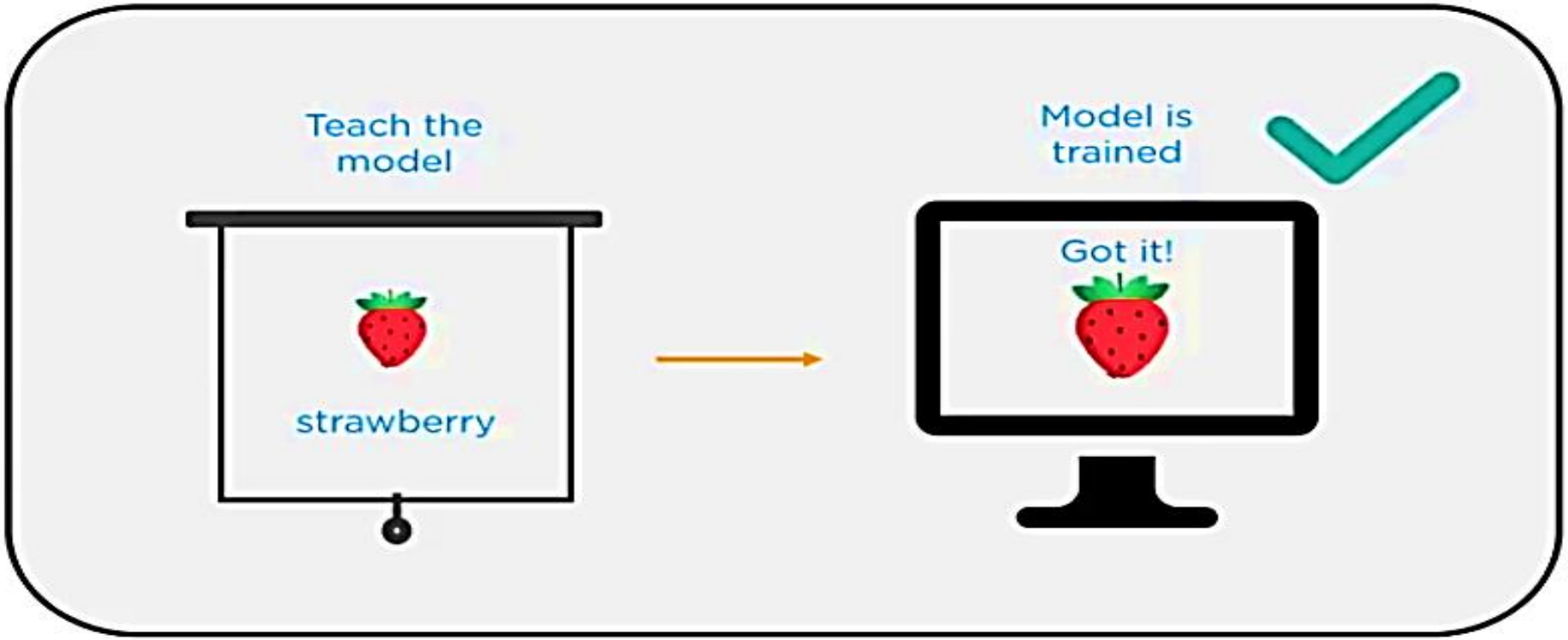
1. Why Support Vector Machine?
2. What is Support Vector Machine?
3. Understanding Support Vector Machine

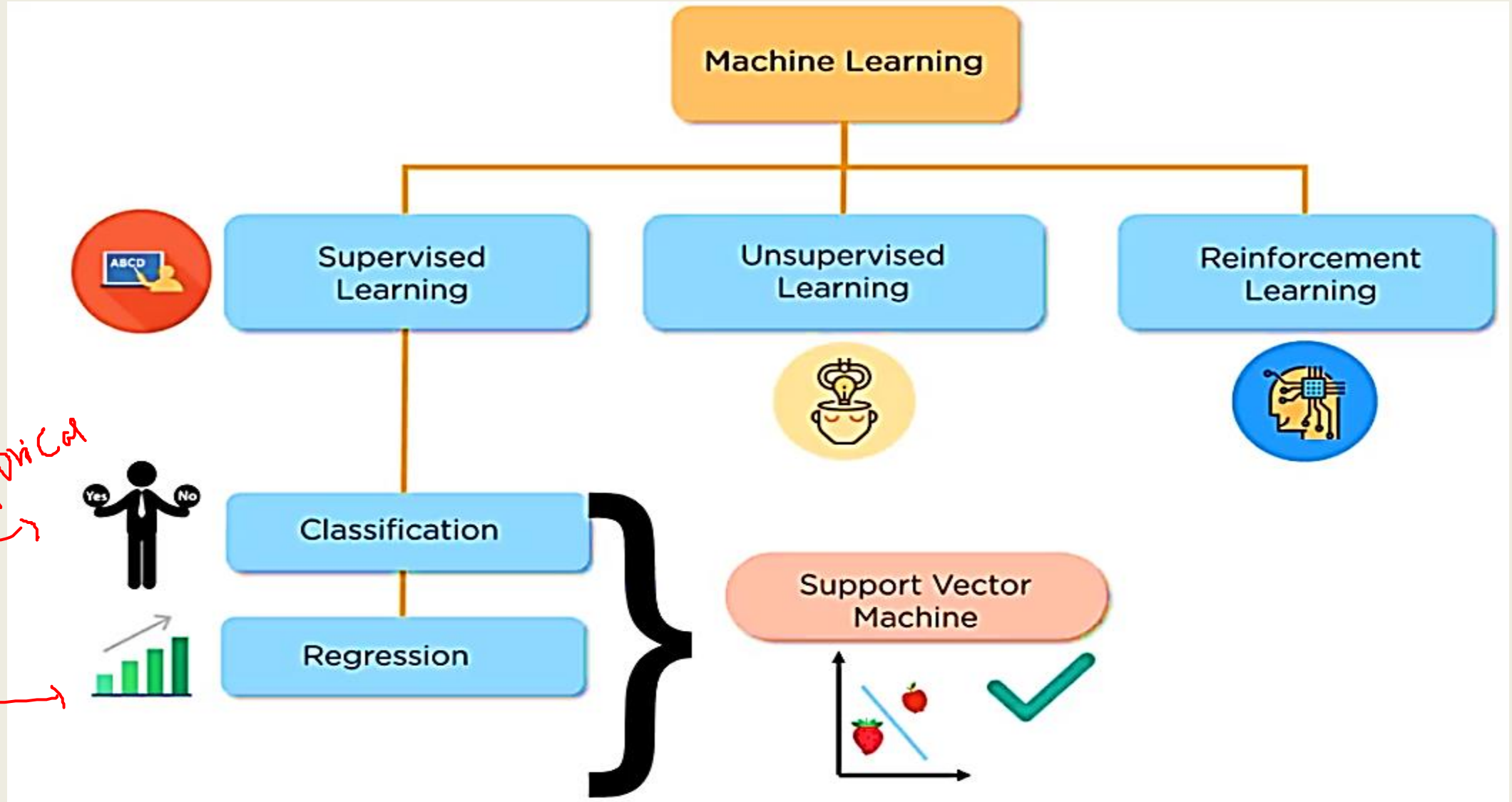
# What is Machine Learning?



## Supervised Learning

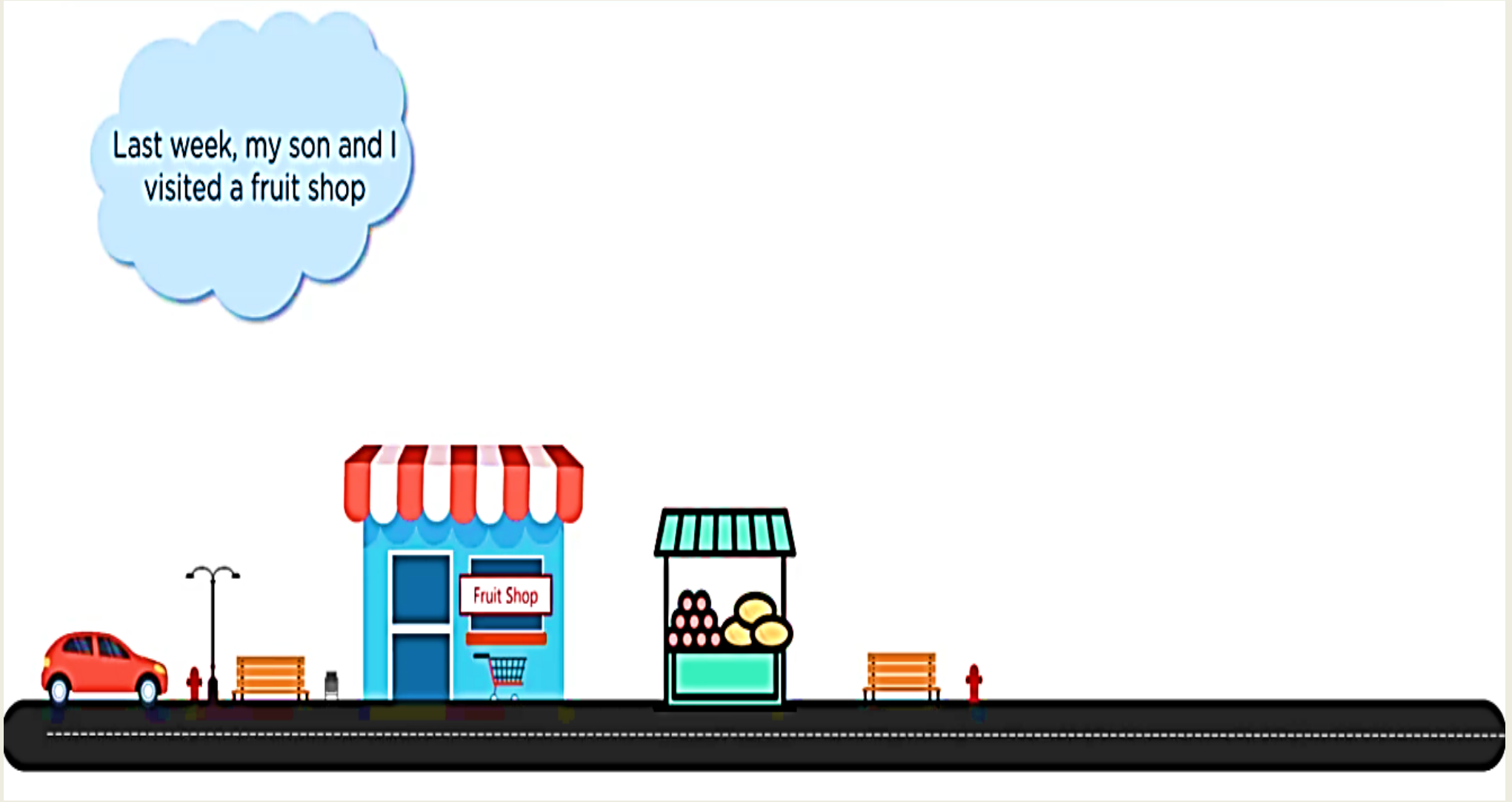
Machine learning model learns from the past input data and makes future prediction as output



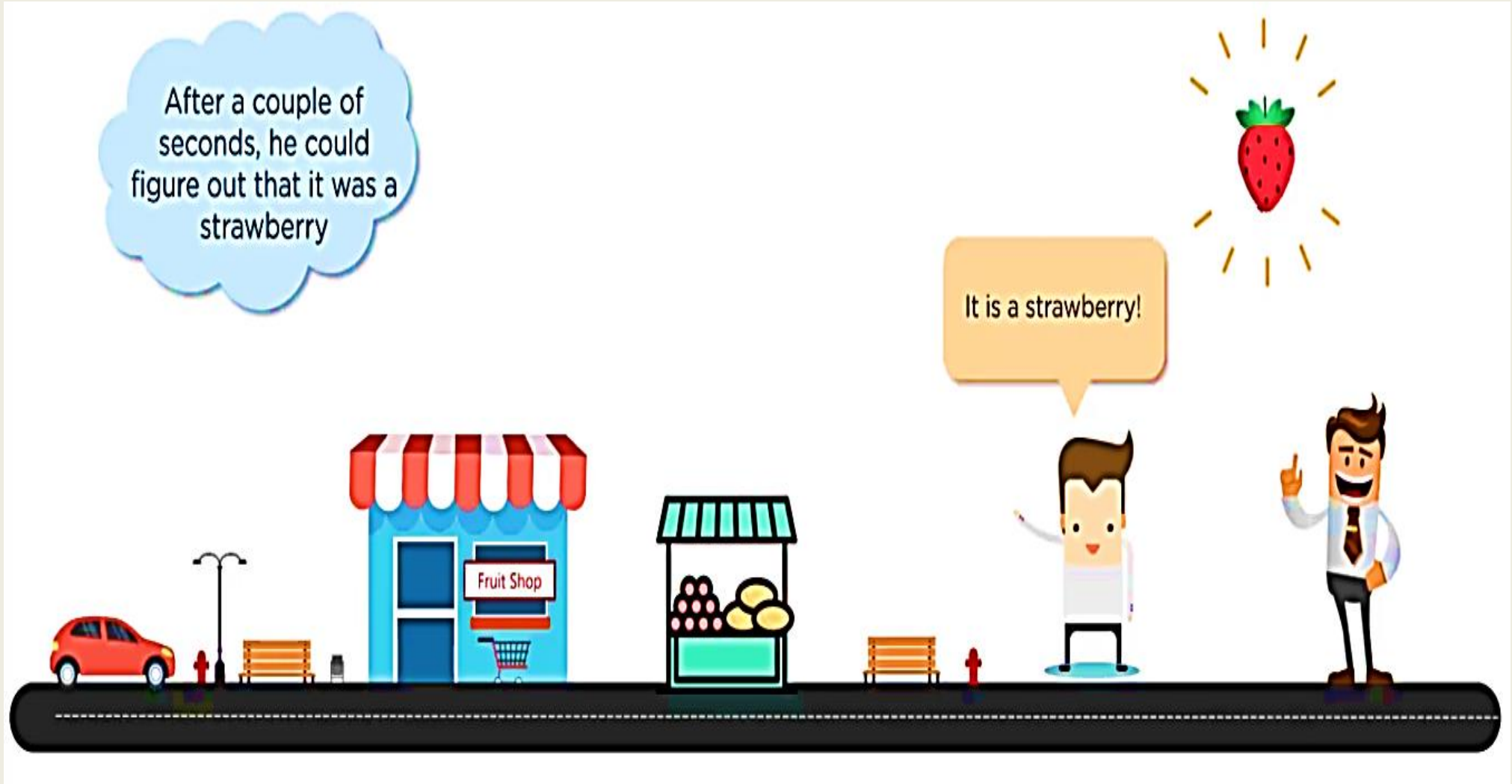

Machine Learning
Supervised Learning
Unsupervised Learning
Reinforcement Learning
ABCD
Yes
No
Classification
Regression
Support Vector Machine
Categorical
Continuous

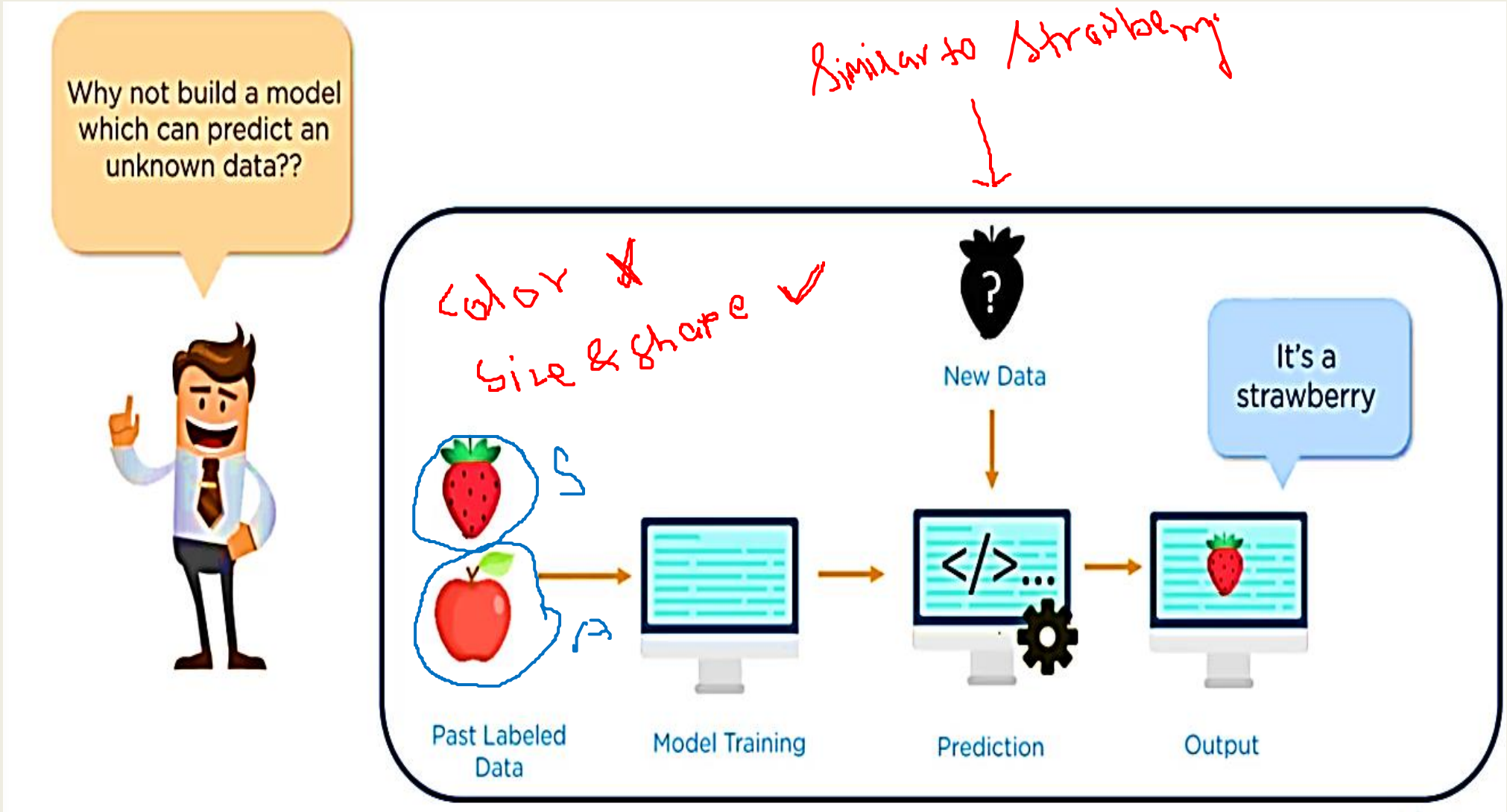
# Why Support Vector Machine?

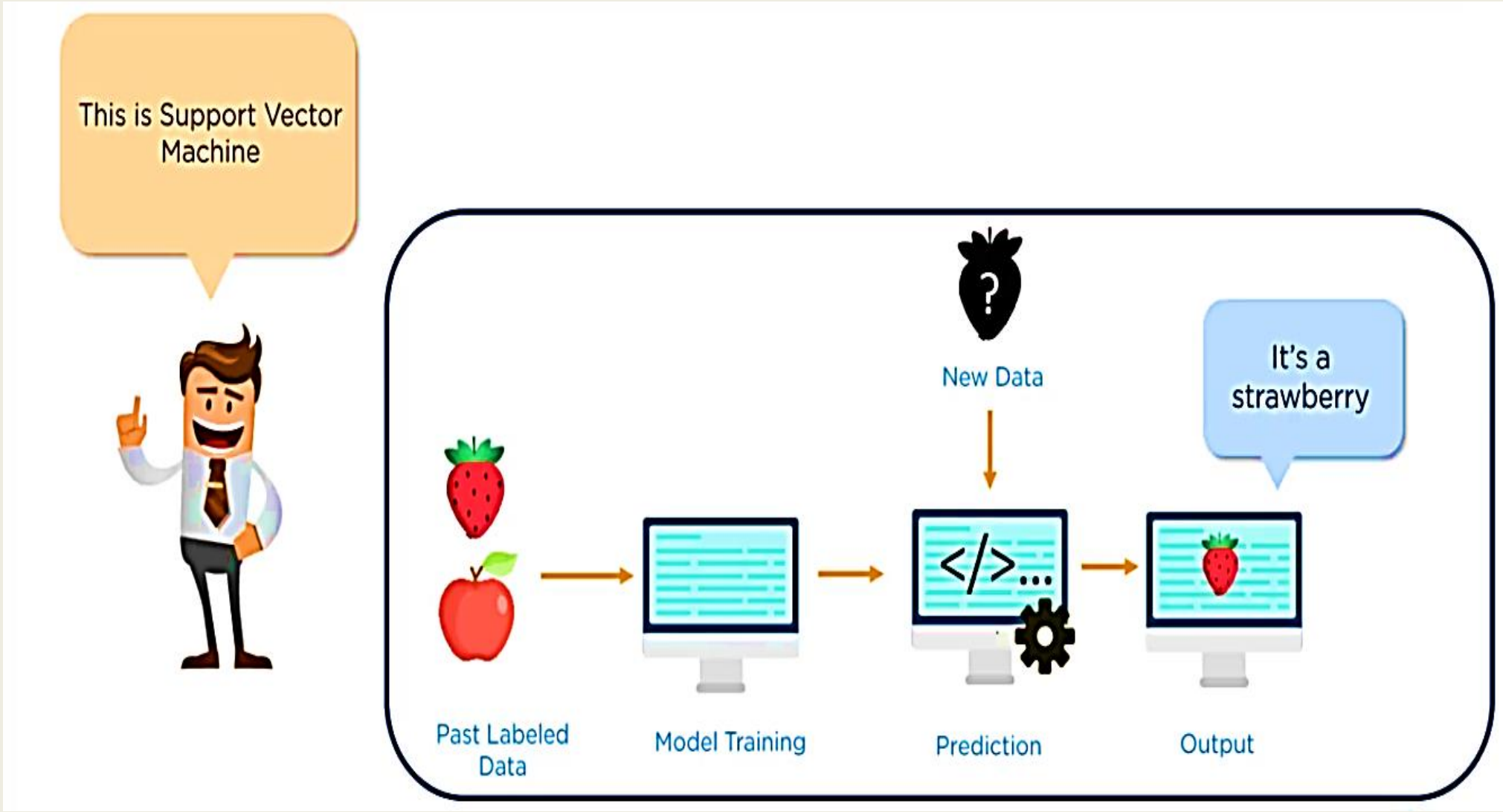
(SVM)

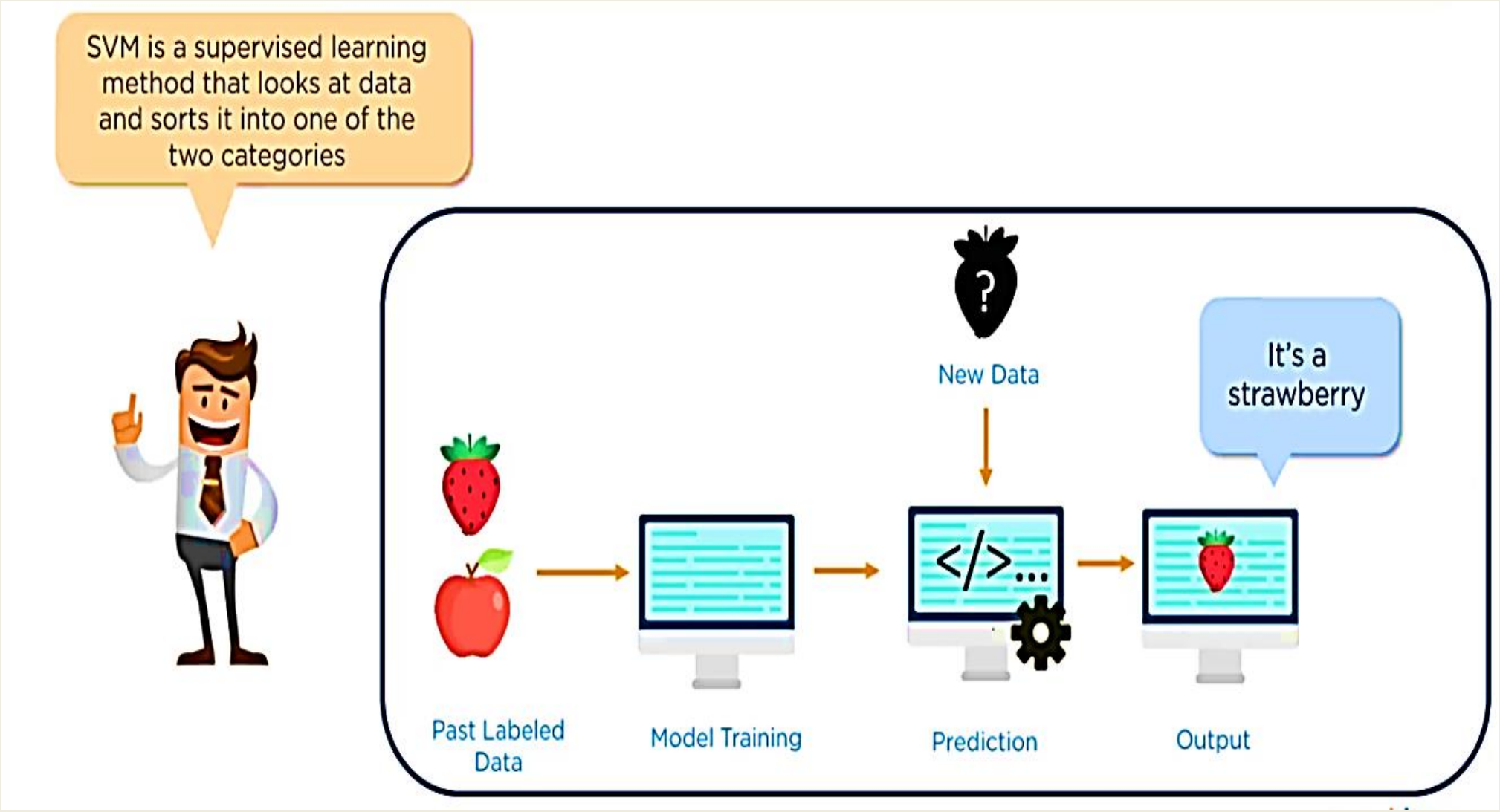

Last week, my son and I visited a fruit shop
Fruit Shop

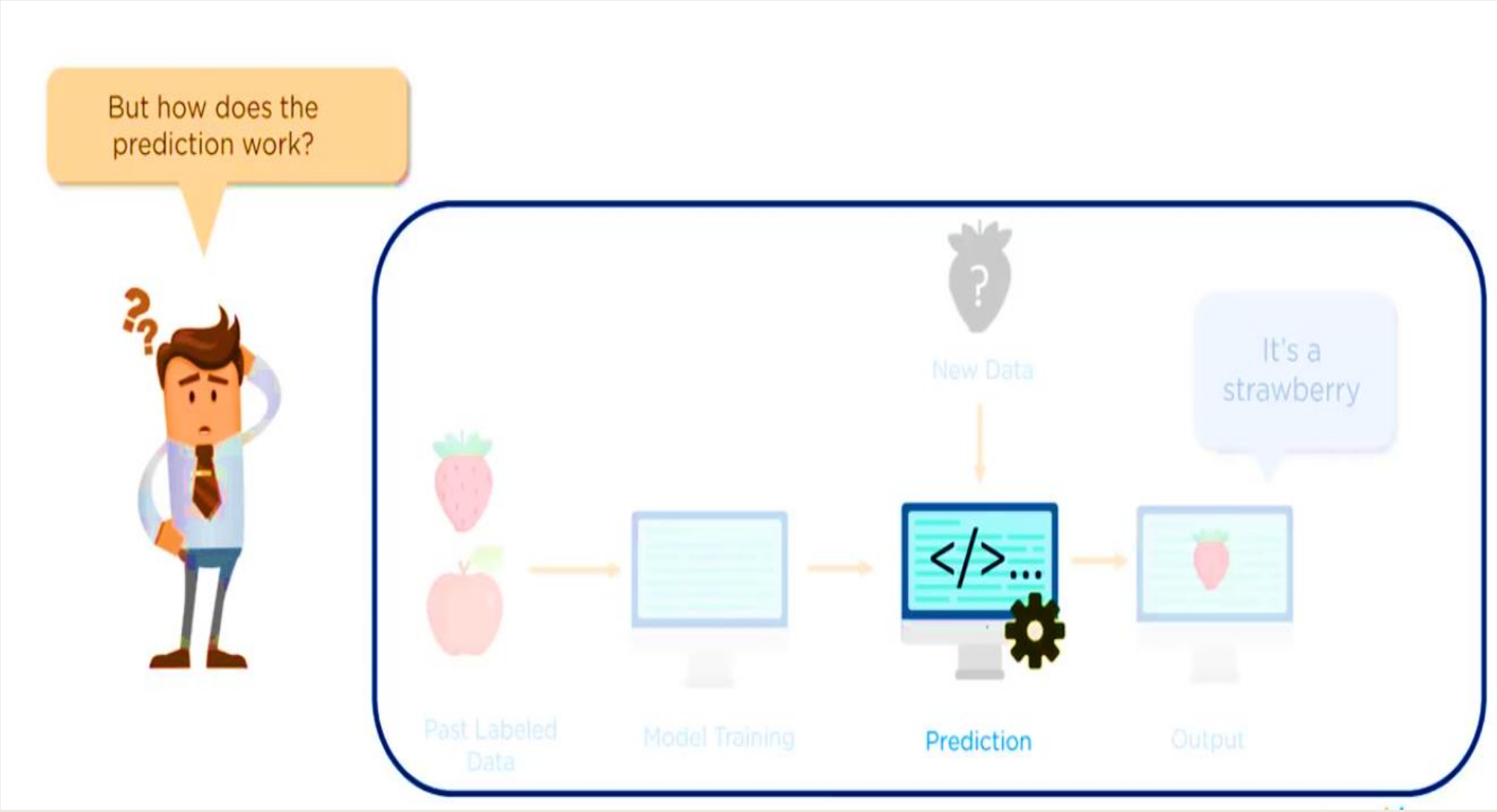

There, he found a fruit which was similar to both apple and strawberry
Dad, is that an apple or a strawberry?
?
Fruit Shop

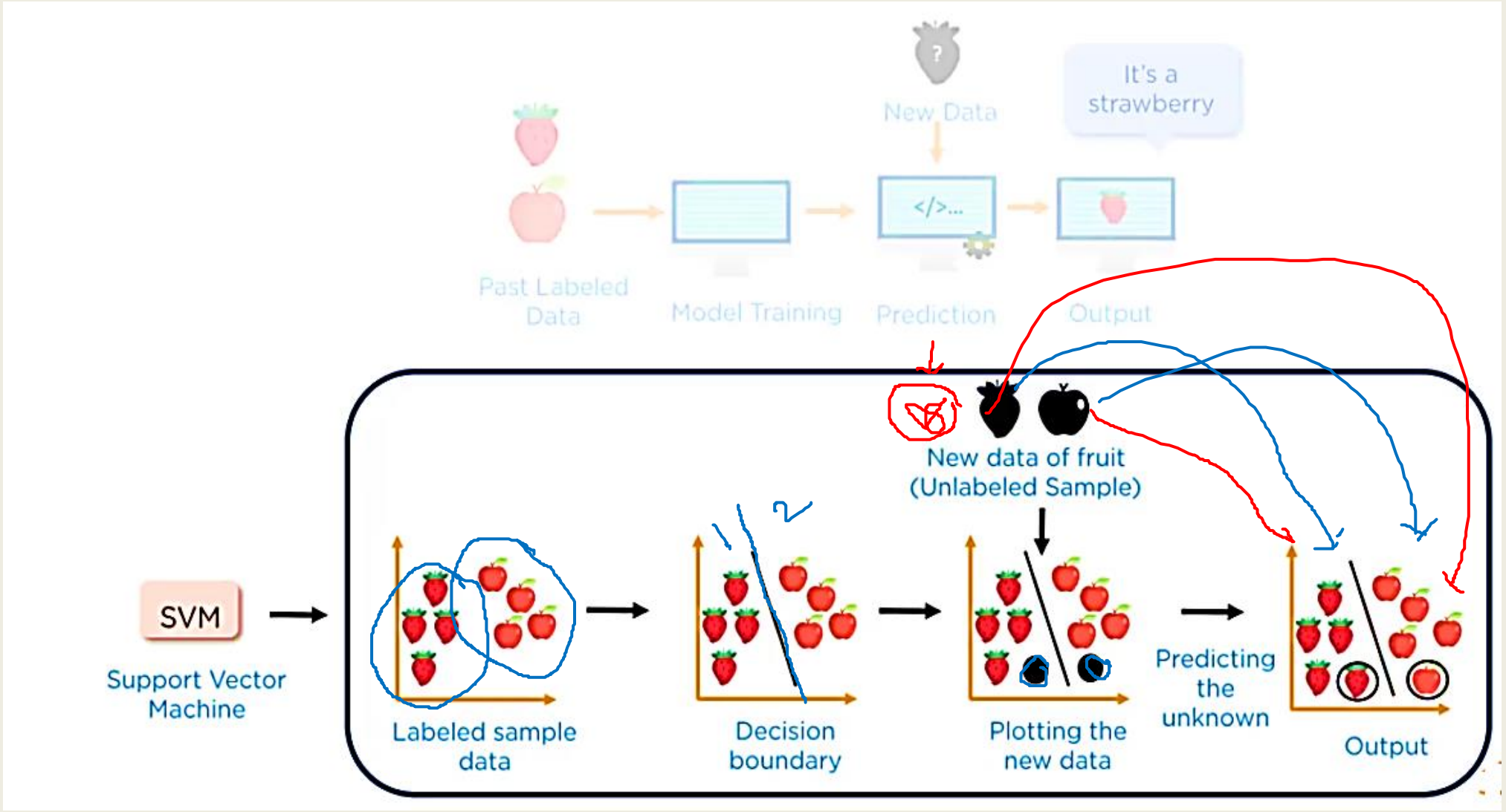

After a couple of seconds, he could figure out that it was a strawberry
It is a strawberry!
Fruit Shop


Why not build a model which can predict an unknown data??
Similar to Strawberry
Color ✗
Size & Shape ✓
S
A
New Data
It's a strawberry
Past Labeled Data
Model Training
Prediction
Output


This is Support Vector Machine
New Data
It's a strawberry
Past Labeled Data
Model Training
Prediction
Output

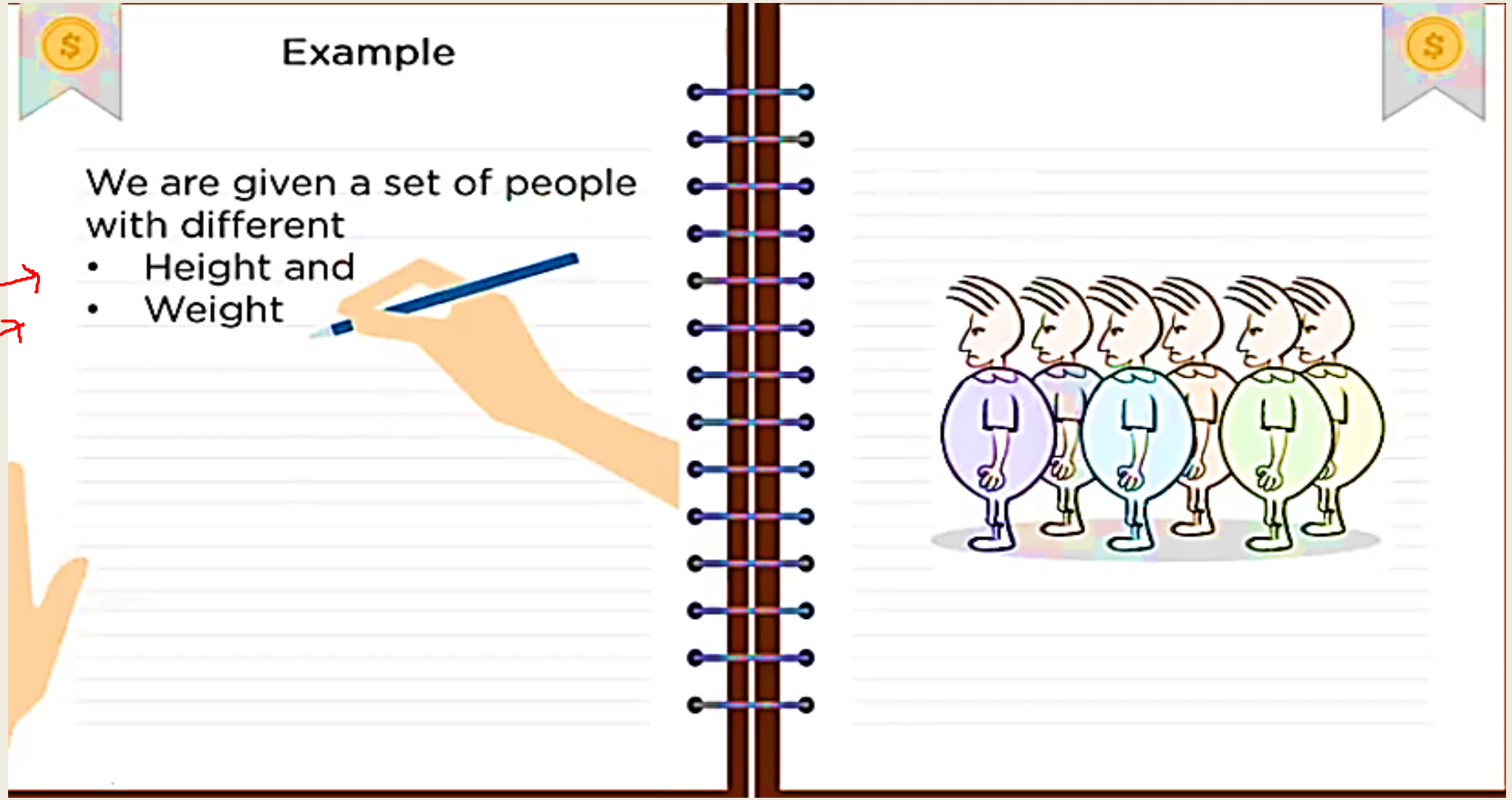

SVM is a supervised learning method that looks at data and sorts it into one of the two categories
New Data
It's a strawberry
Past Labeled Data
Model Training
Prediction
Output


But how does the prediction work?
New Data
It's a strawberry
Past Labeled Data
Model Training
Prediction
Output

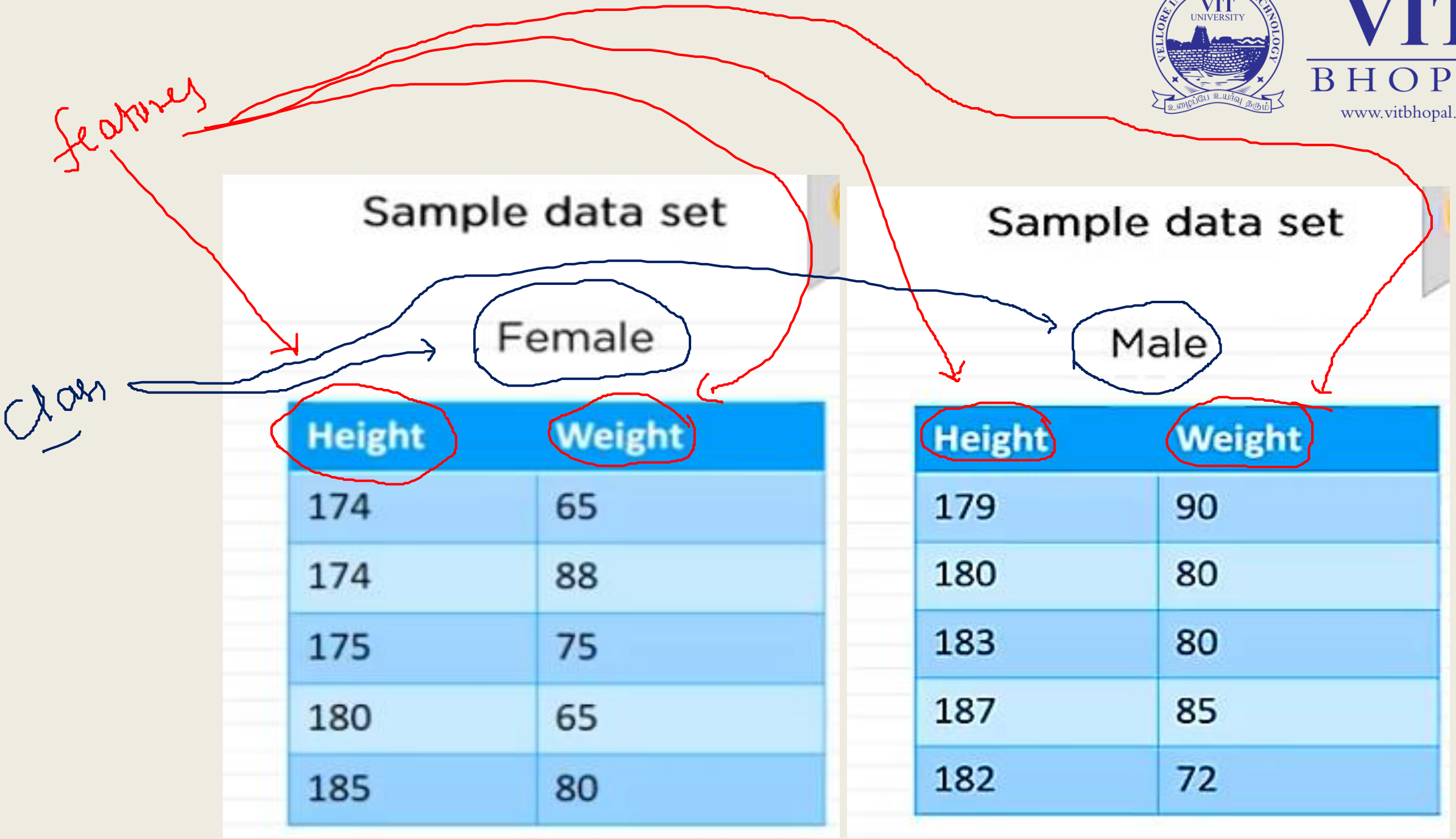

New Data
It's a strawberry
Past Labeled Data
Model Training
Prediction
Output
New data of fruit (Unlabeled Sample)
SVM
Support Vector Machine
Labeled sample data
Decision boundary
Plotting the new data
Predicting the unknown
Output

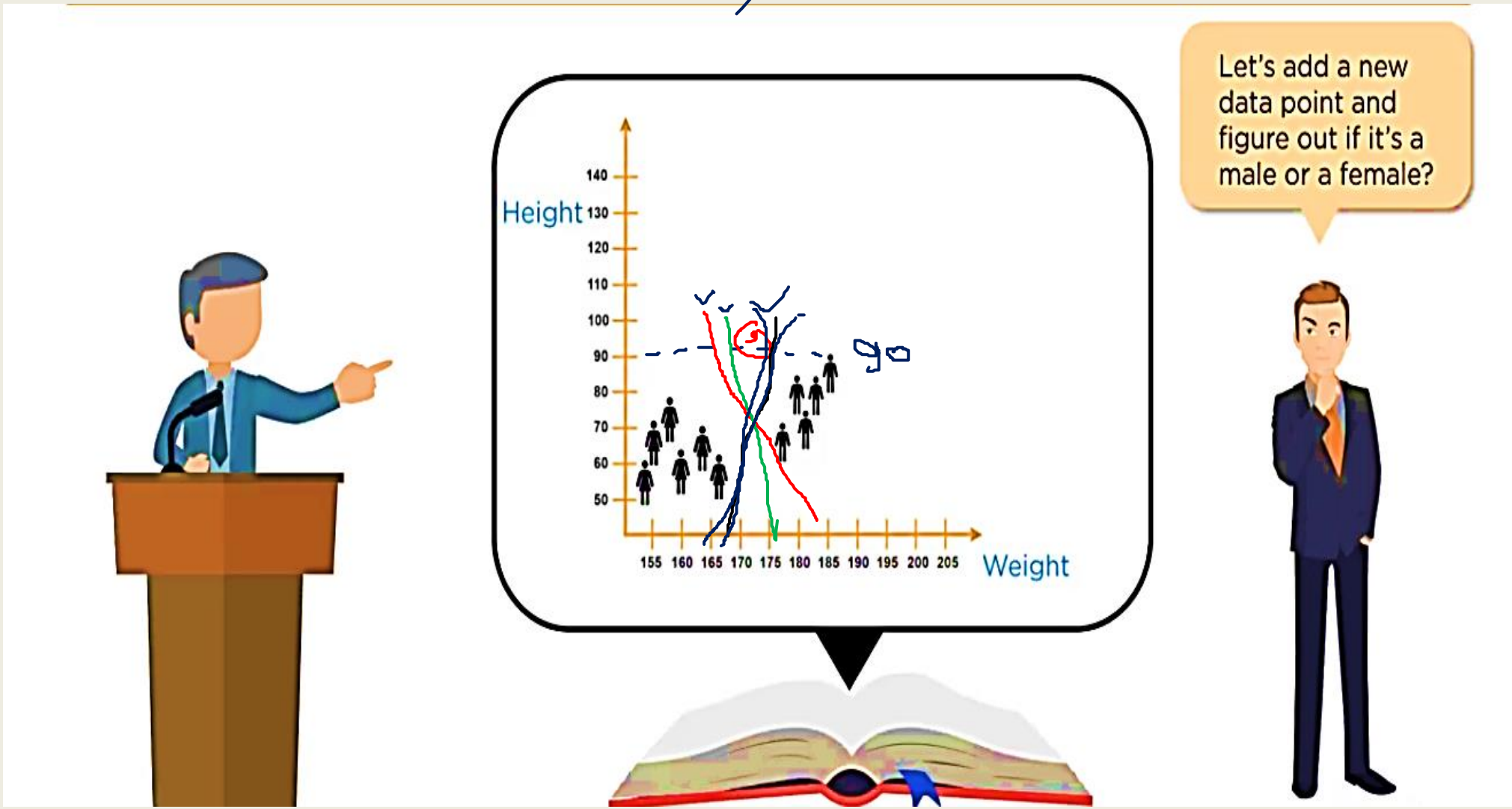
# What is Support Vector Machine?

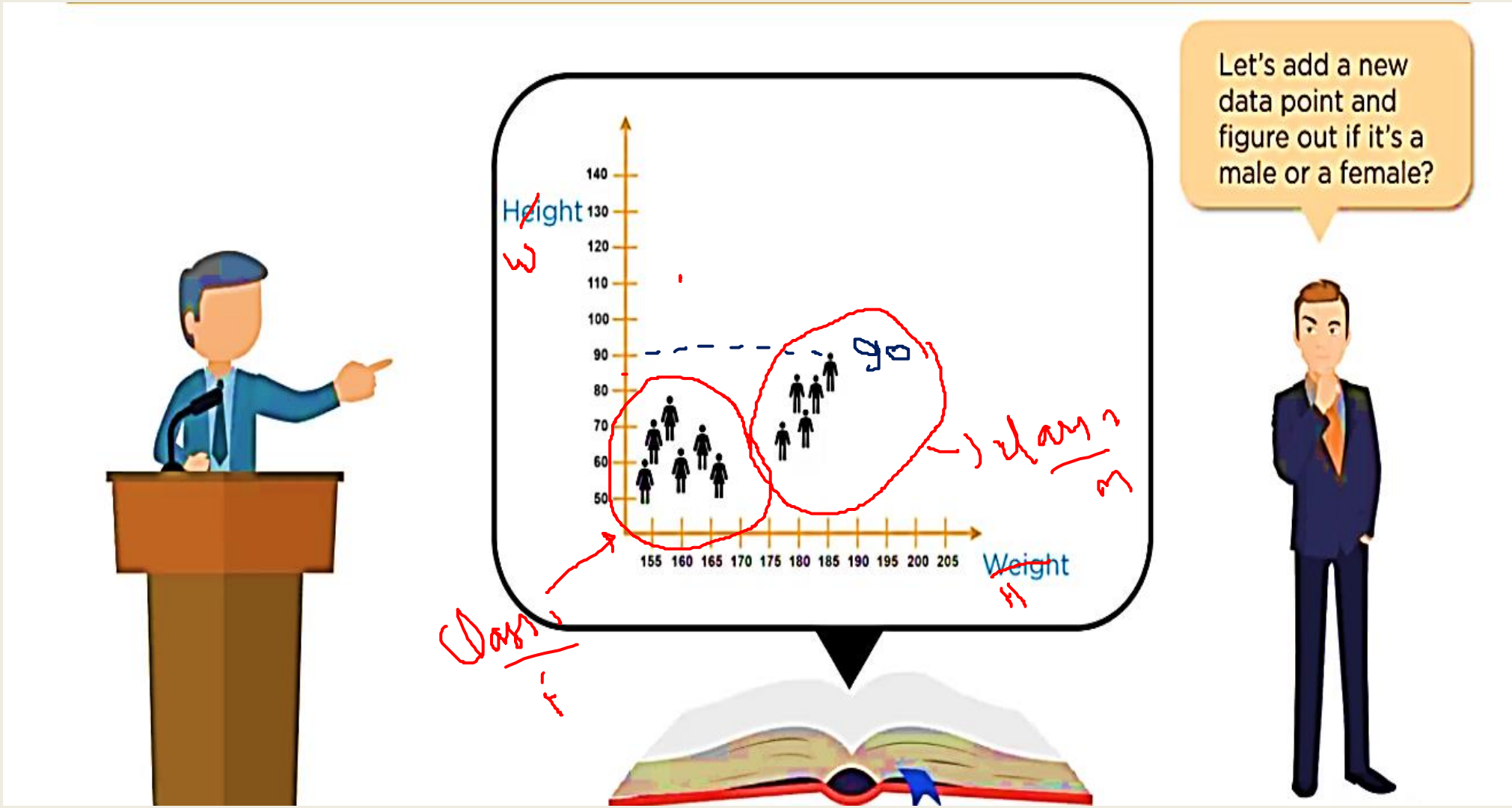
# Example



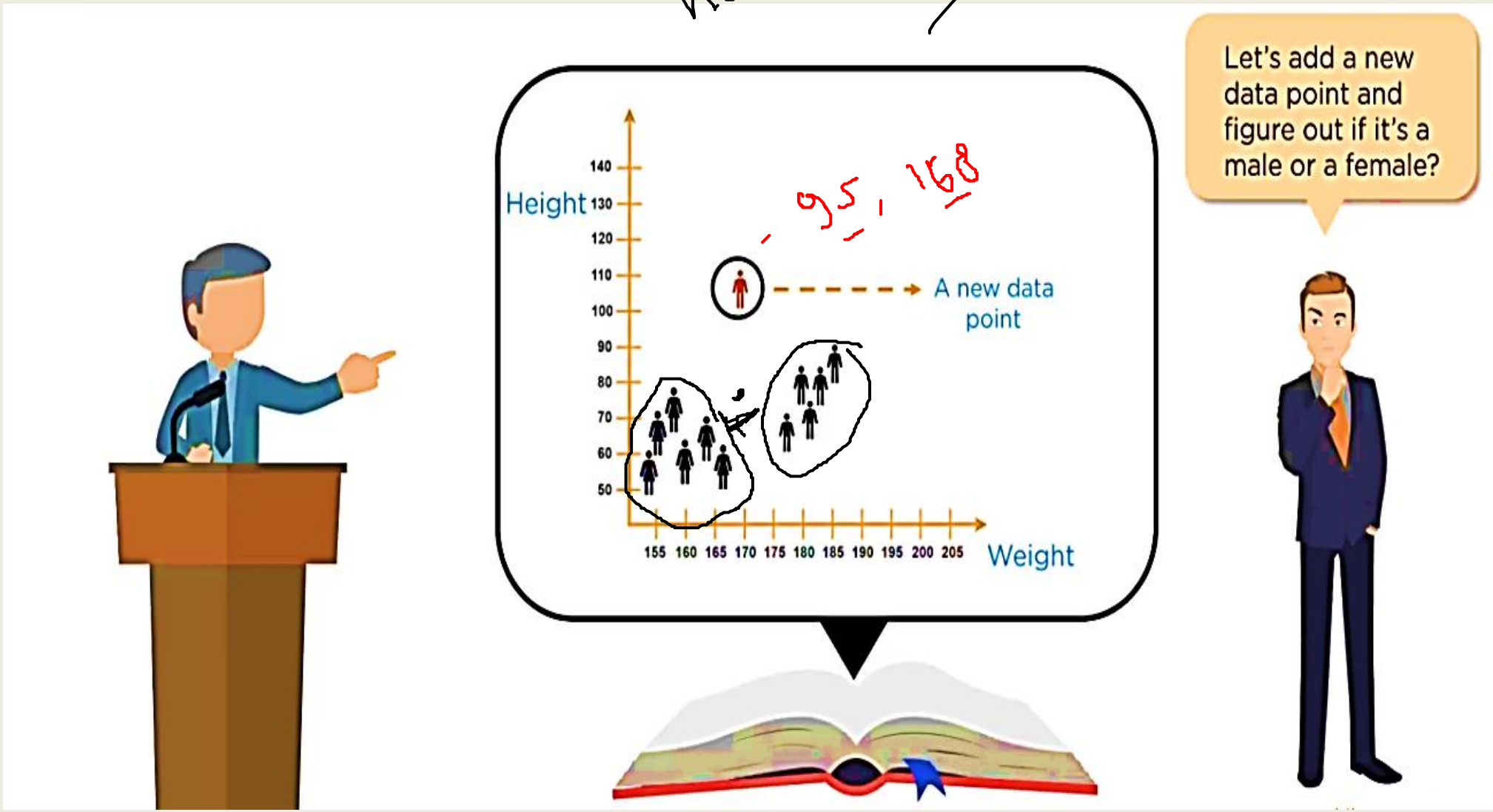

features

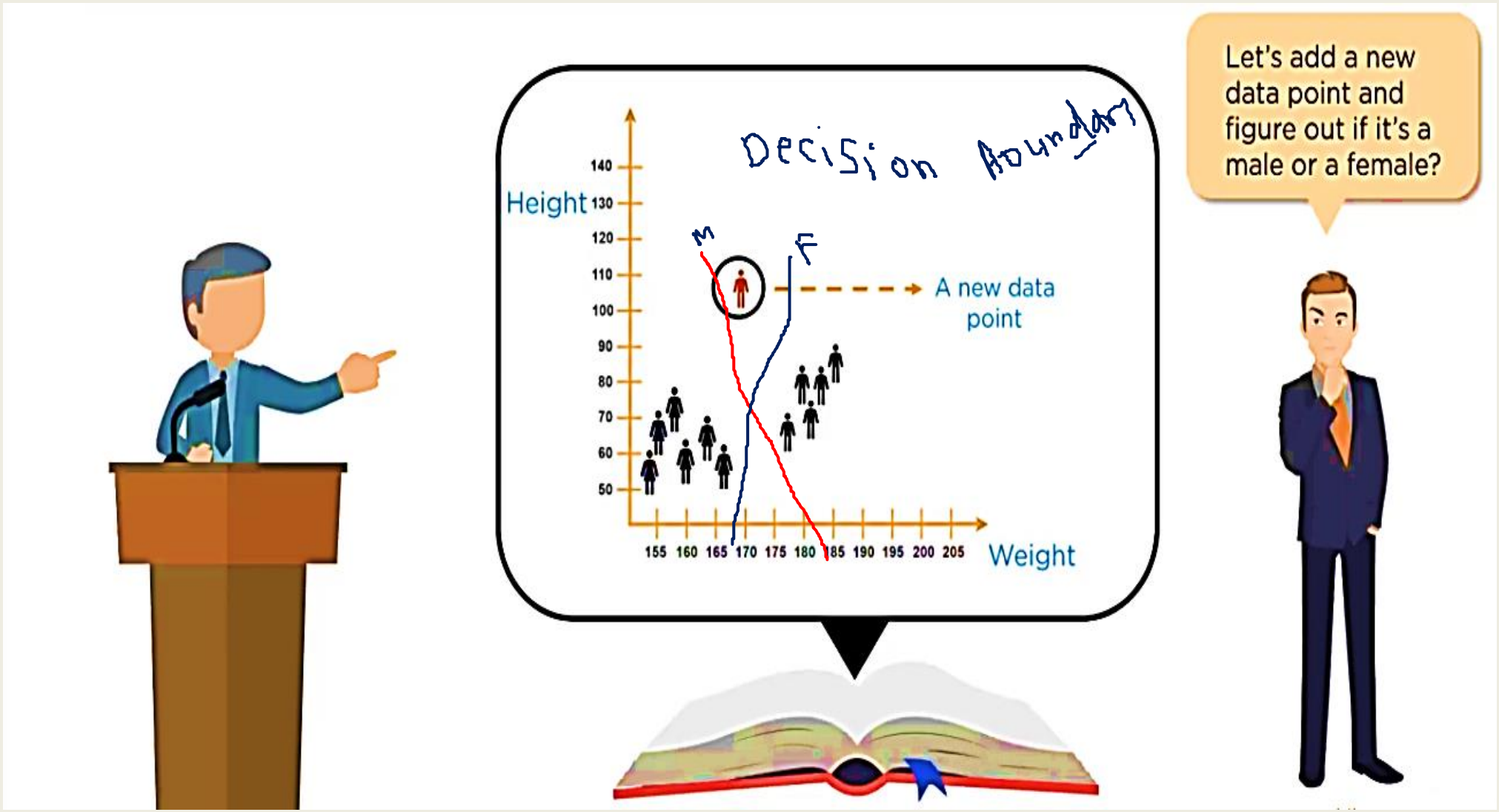

features
class
Sample data set
Female
Height
Weight
174
65
174
88
175
75
180
65
185
80
Sample data set
Male
Height
Weight
179
90
180
80
183
80
187
85
182
72

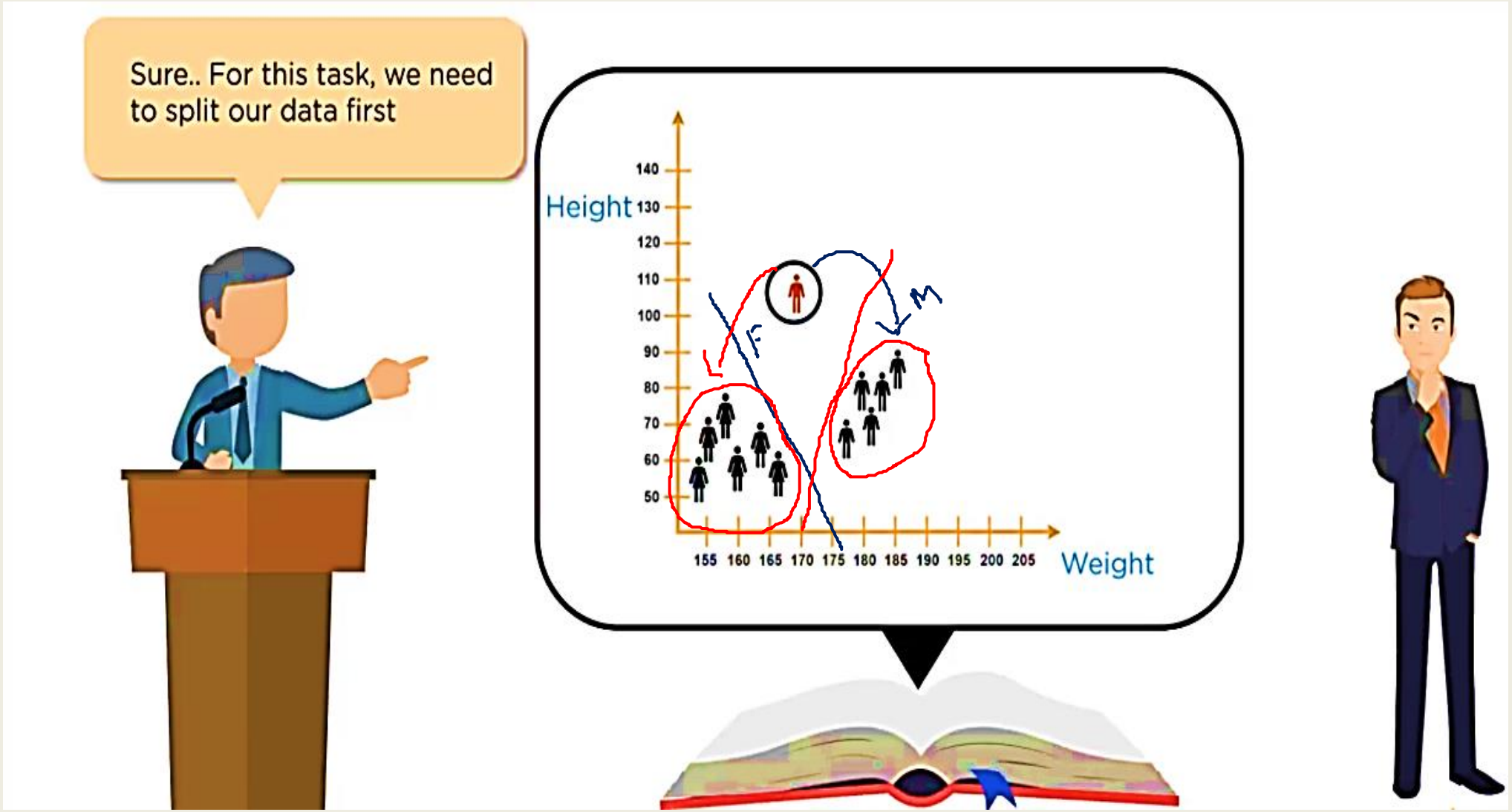

Let's add a new data point and figure out if it's a male or a female?
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight

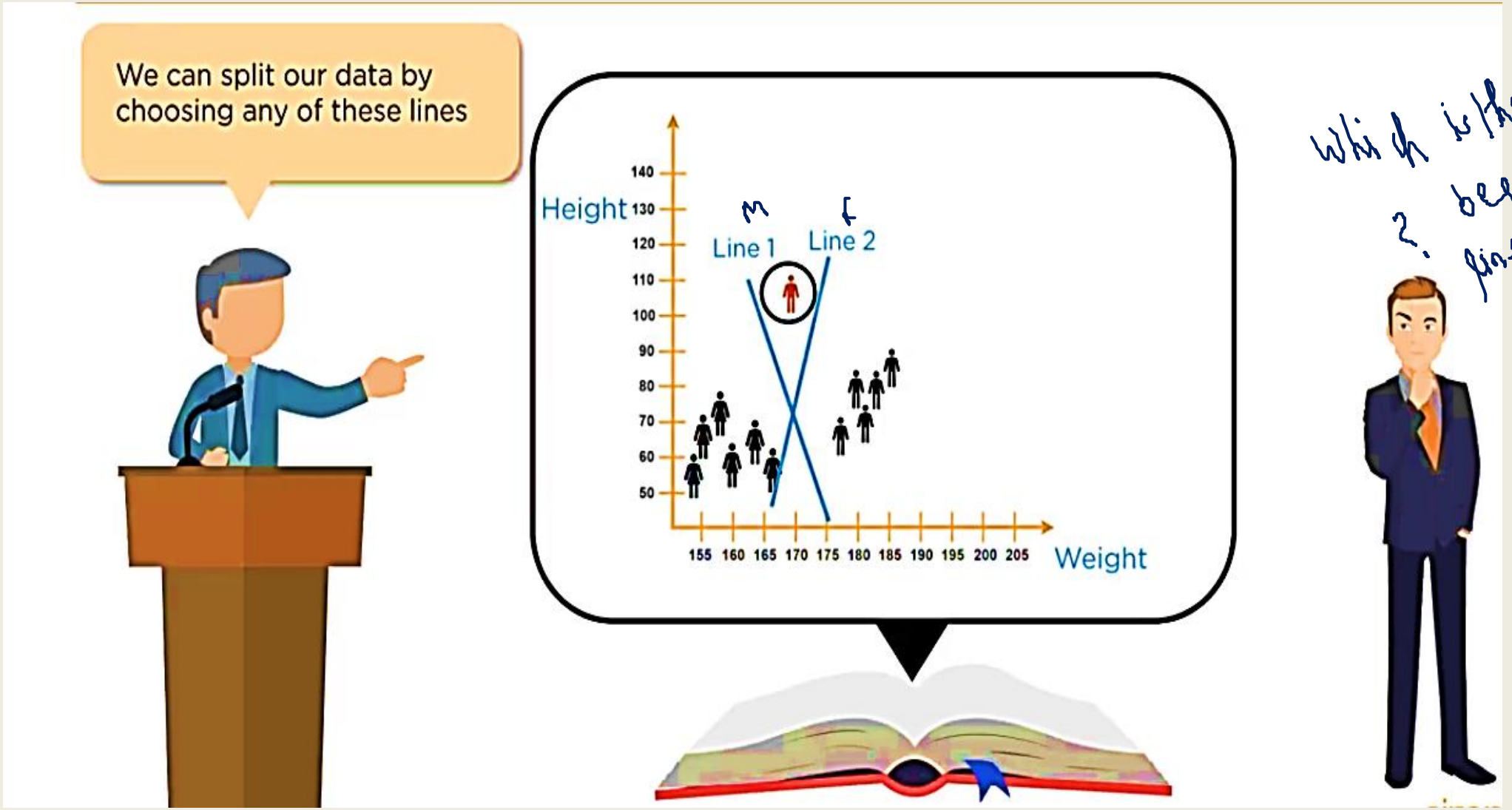

Let's add a new data point and figure out if it's a male or a female?
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight

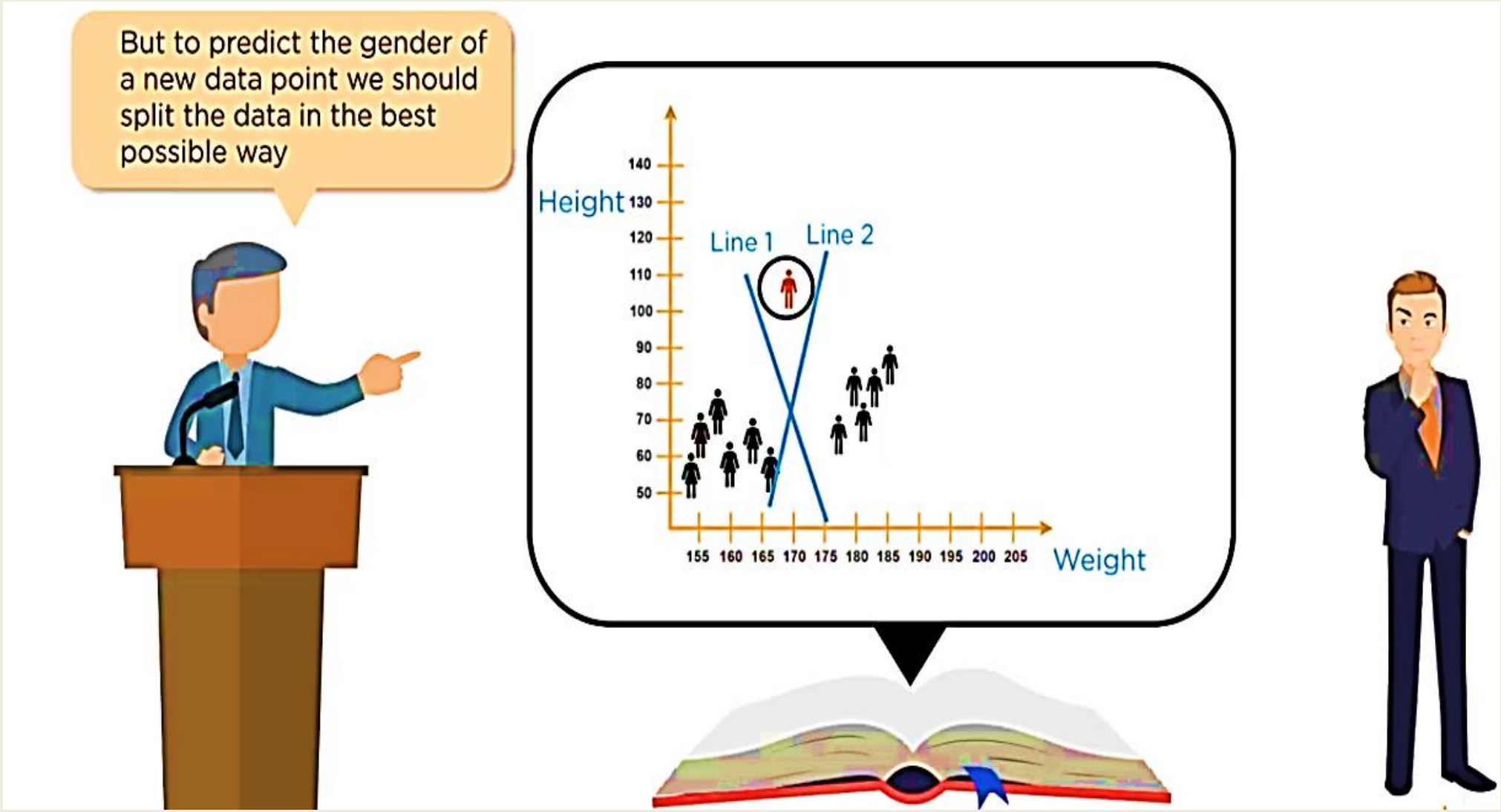
n- line ($y = mx + c$)

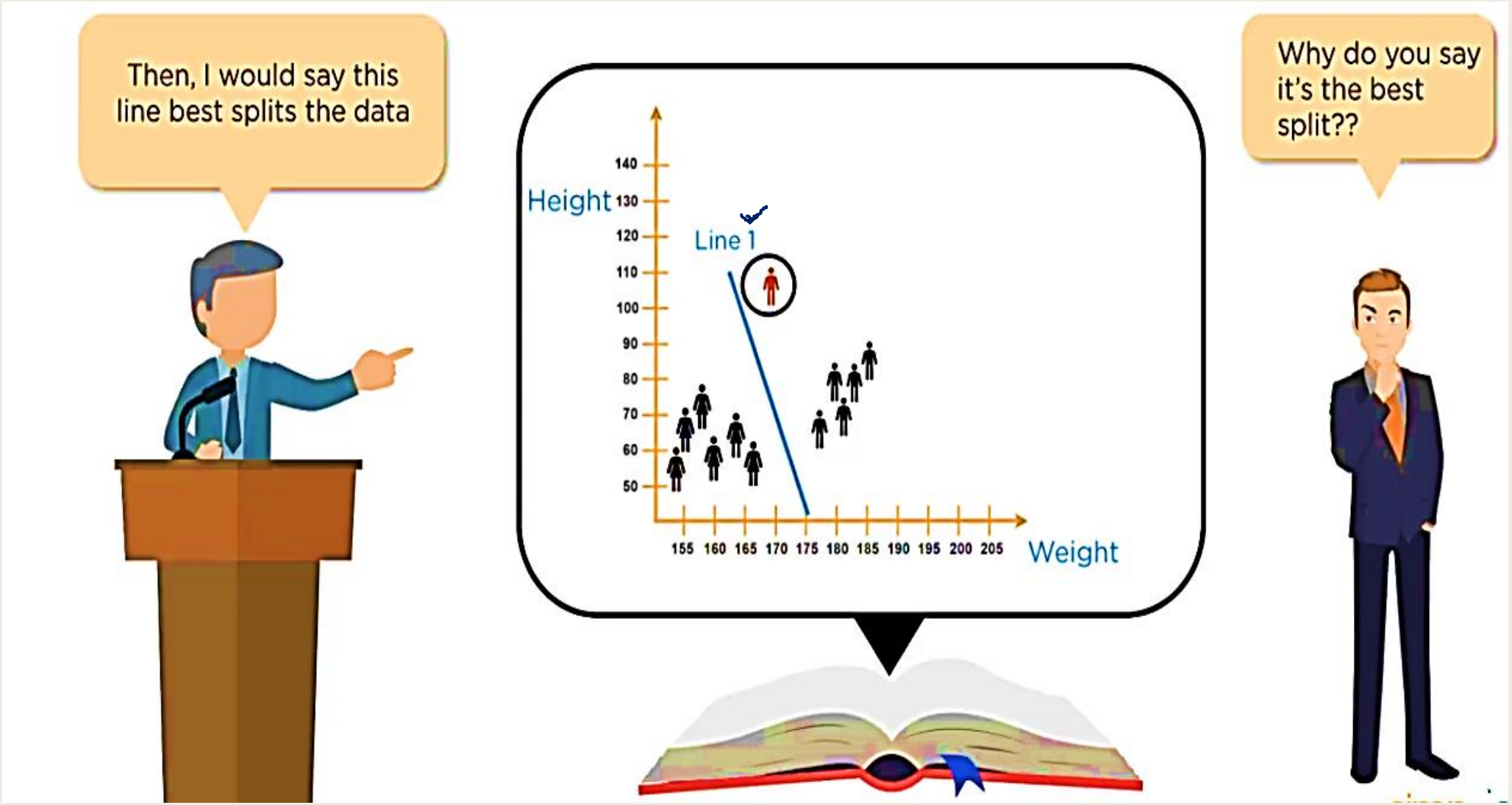

Let's add a new data point and figure out if it's a male or a female?
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
A new data point
95, 168

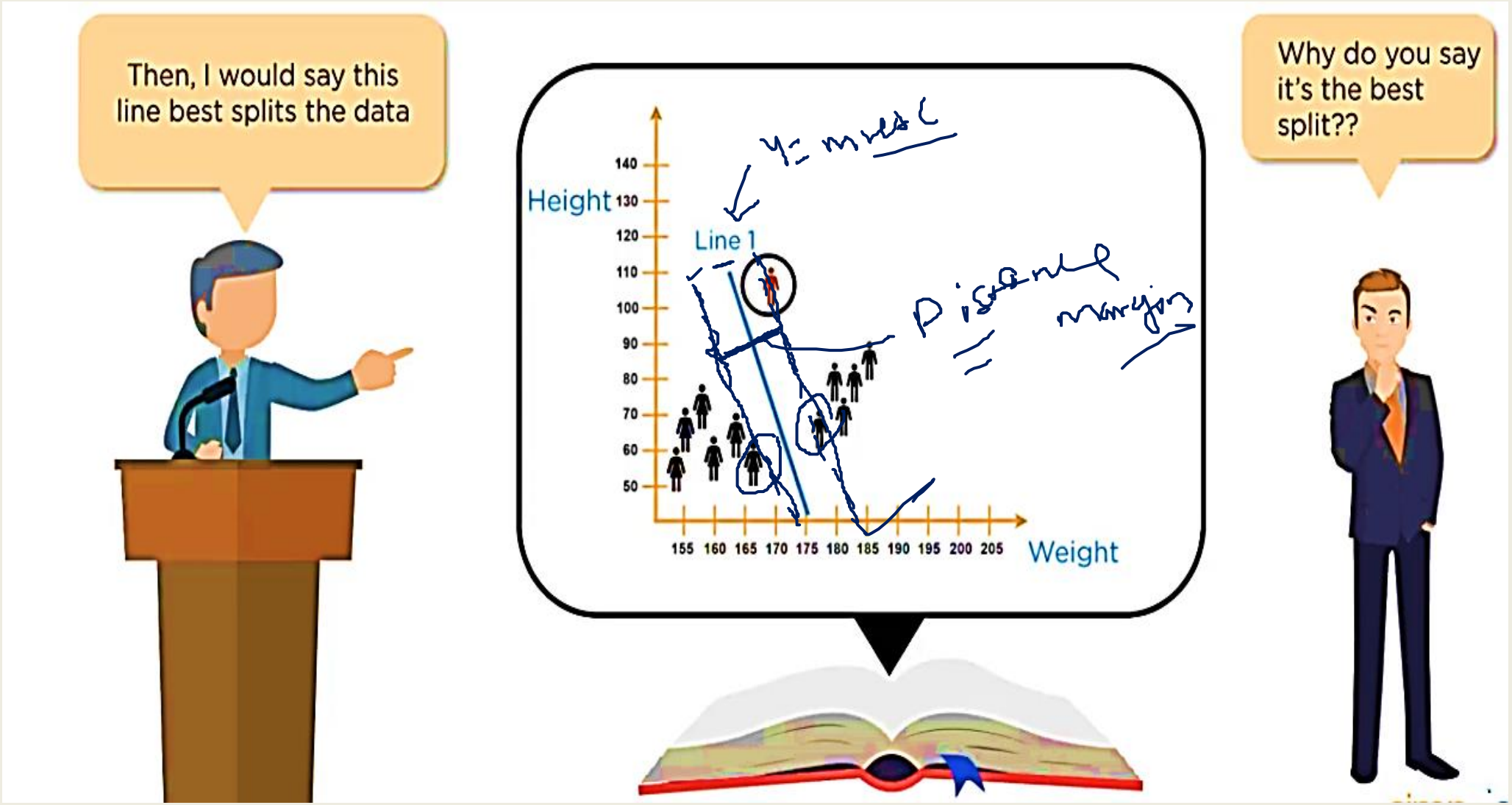

Let's add a new data point and figure out if it's a male or a female?
Decision Boundary
Height
140
130
120
110
100
90
80
70
60
50
M
F
A new data point
155 160 165 170 175 180 185 190 195 200 205
Weight

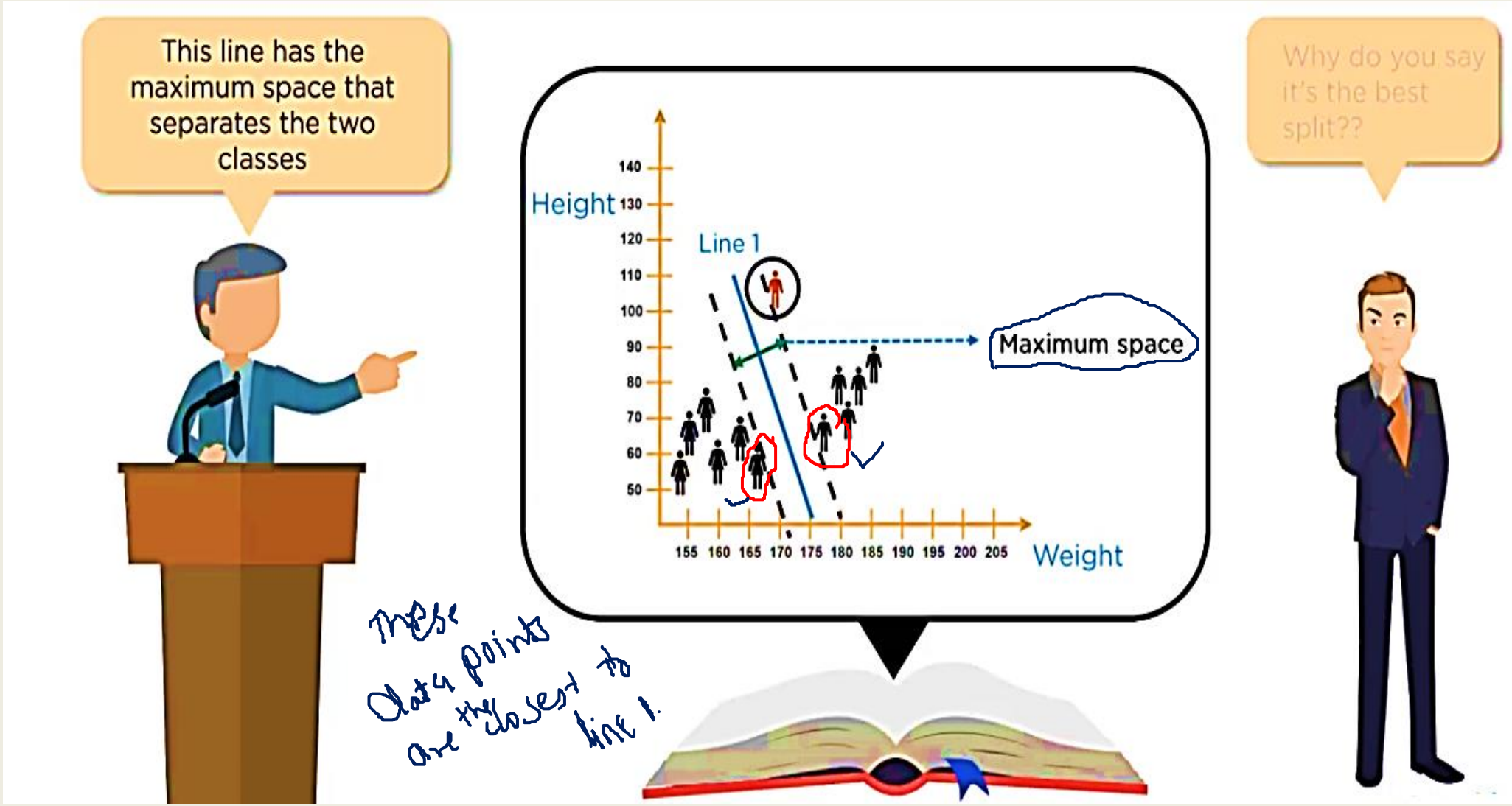

Sure.. For this task, we need to split our data first
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
F
M

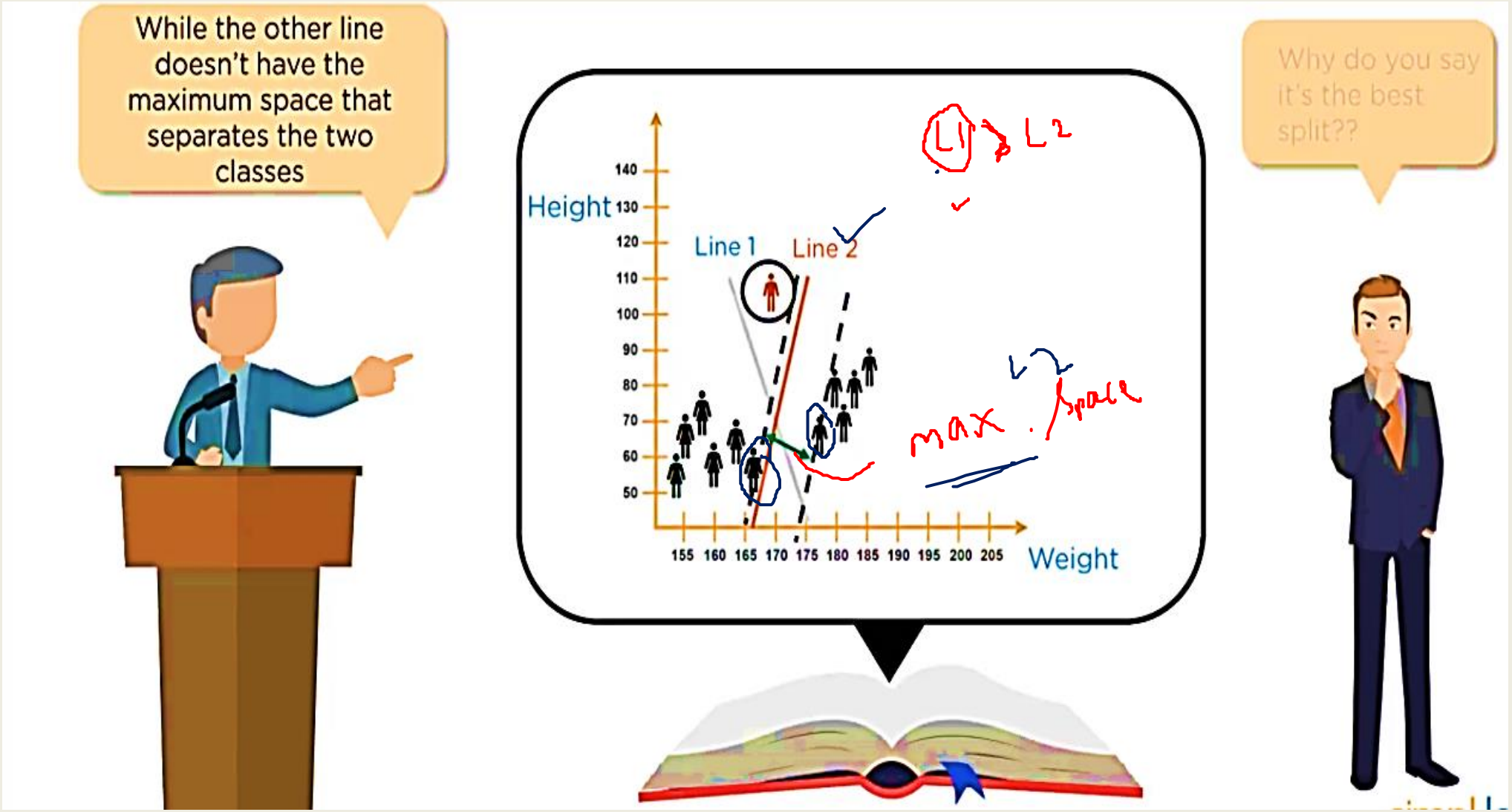

We can split our data by choosing any of these lines
Height
140
130
120
110
100
90
80
70
60
50
m
f
Line 1
Line 2
155 160 165 170 175 180 185 190 195 200 205
Weight
Which is the best line?

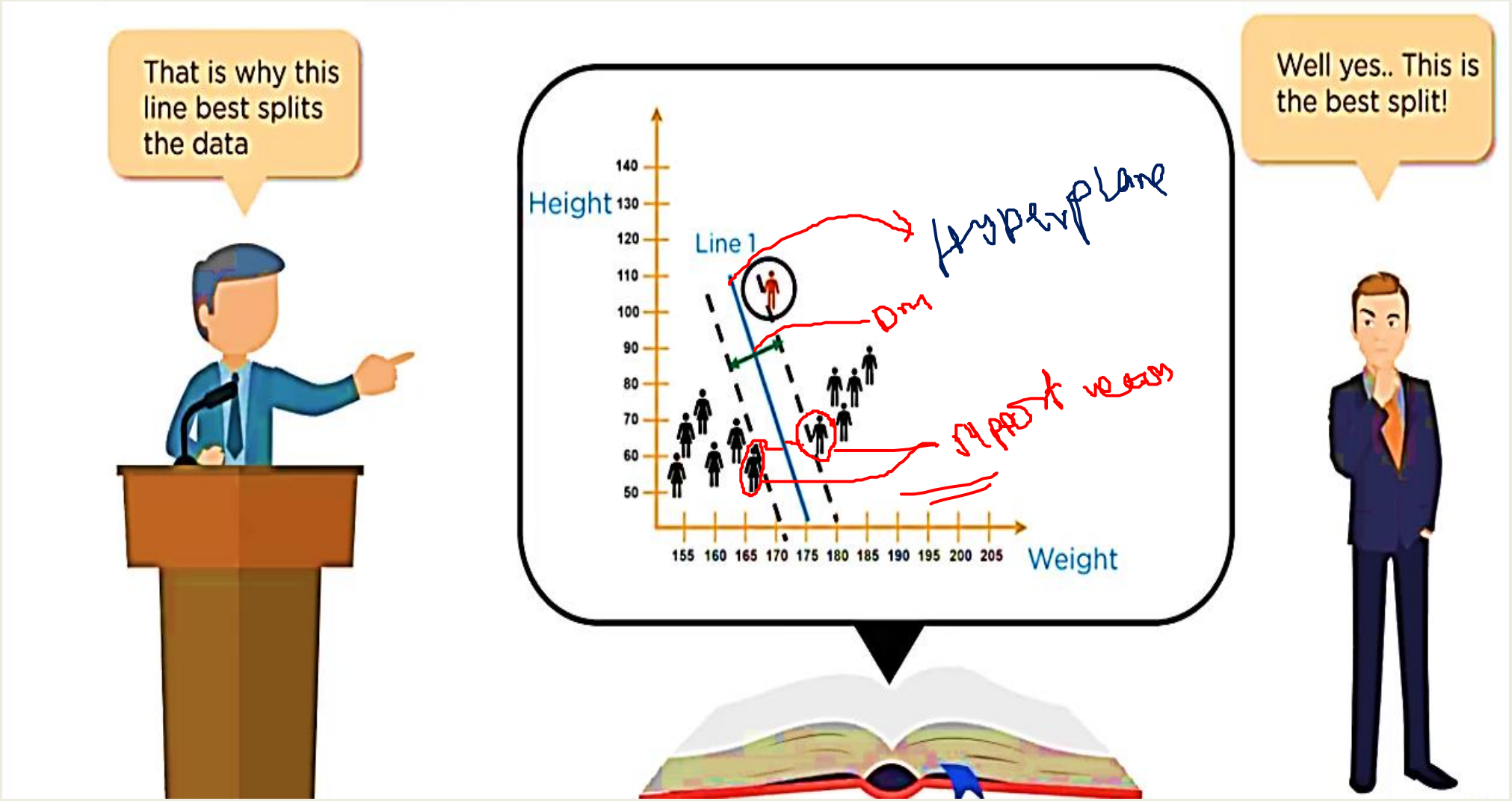

But to predict the gender of a new data point we should split the data in the best possible way
Height
140
130
120
110
100
90
80
70
60
50
Line 1
Line 2
155 160 165 170 175 180 185 190 195 200 205
Weight


Then, I would say this line best splits the data
Height
140
130
120
110
100
90
80
70
60
50
Line 1
155 160 165 170 175 180 185 190 195 200 205
Weight
Why do you say it's the best split??


Then, I would say this line best splits the data
Height
140
130
120
110
100
90
80
70
60
50
Line 1
155 160 165 170 175 180 185 190 195 200 205
Weight
Why do you say it's the best split??


This line has the maximum space that separates the two classes
Why do you say it's the best split??
Height
Weight
Line 1
Maximum space
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
These data points are the closest to line 1.


While the other line doesn't have the maximum space that separates the two classes
Height
140
130
120
110
100
90
80
70
60
50
Line 1
Line 2
155 160 165 170 175 180 185 190 195 200 205
Weight
Why do you say it's the best split??

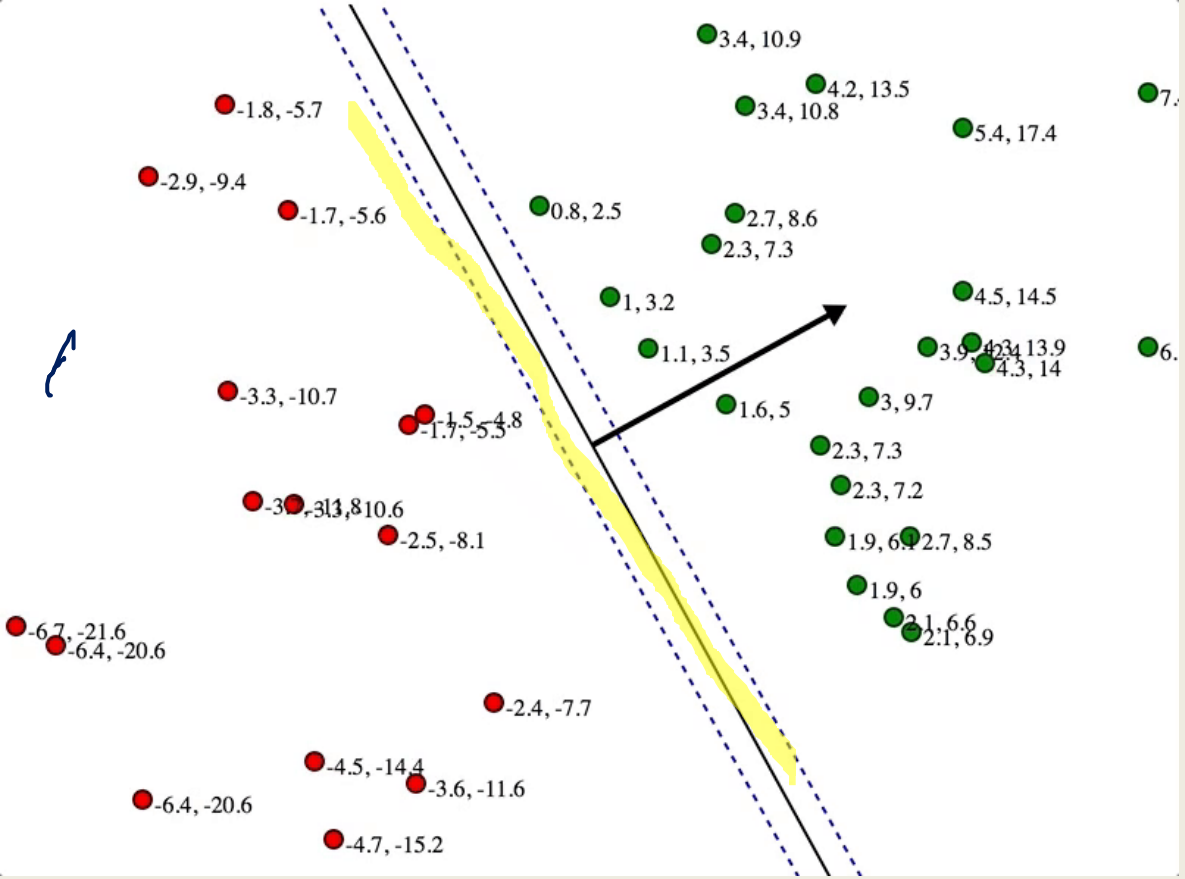

That is why this line best splits the data
Height
140
130
120
110
100
90
80
70
60
50
Line 1
155 160 165 170 175 180 185 190 195 200 205
Weight
Well yes.. This is the best split!

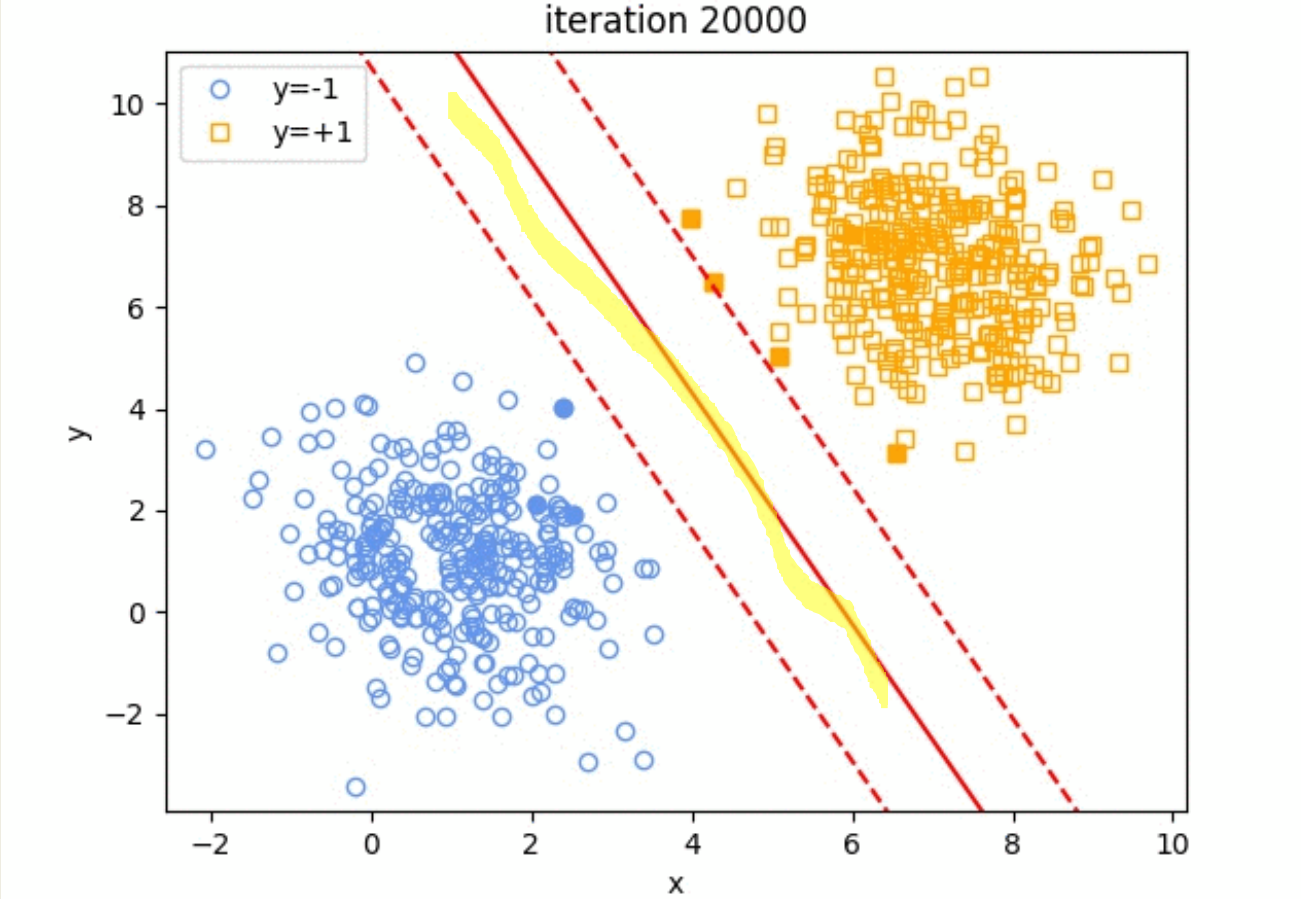
- The followings are important concepts in SVM −

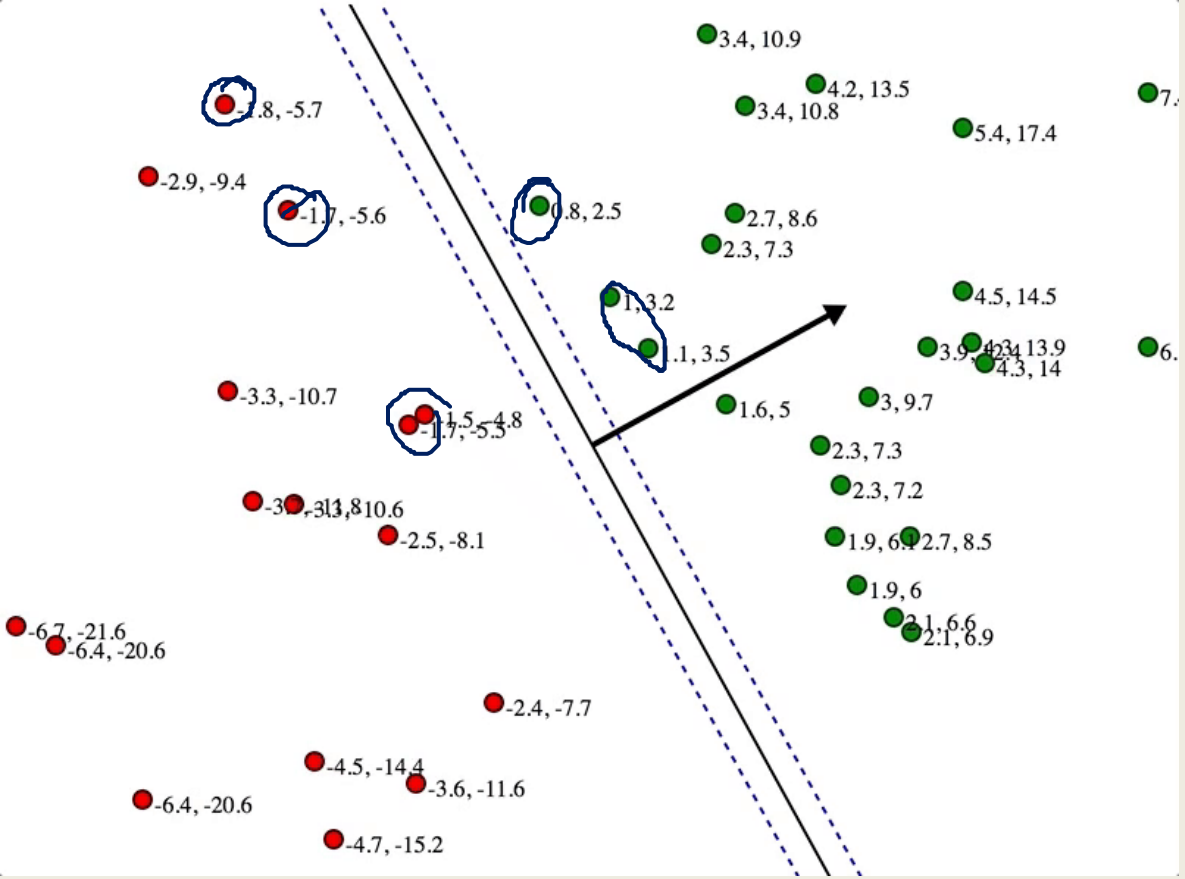
1. Support Vectors − Datapoints that are closest to the hyperplane is called support vectors. Separating line will be defined with the help of these data points.

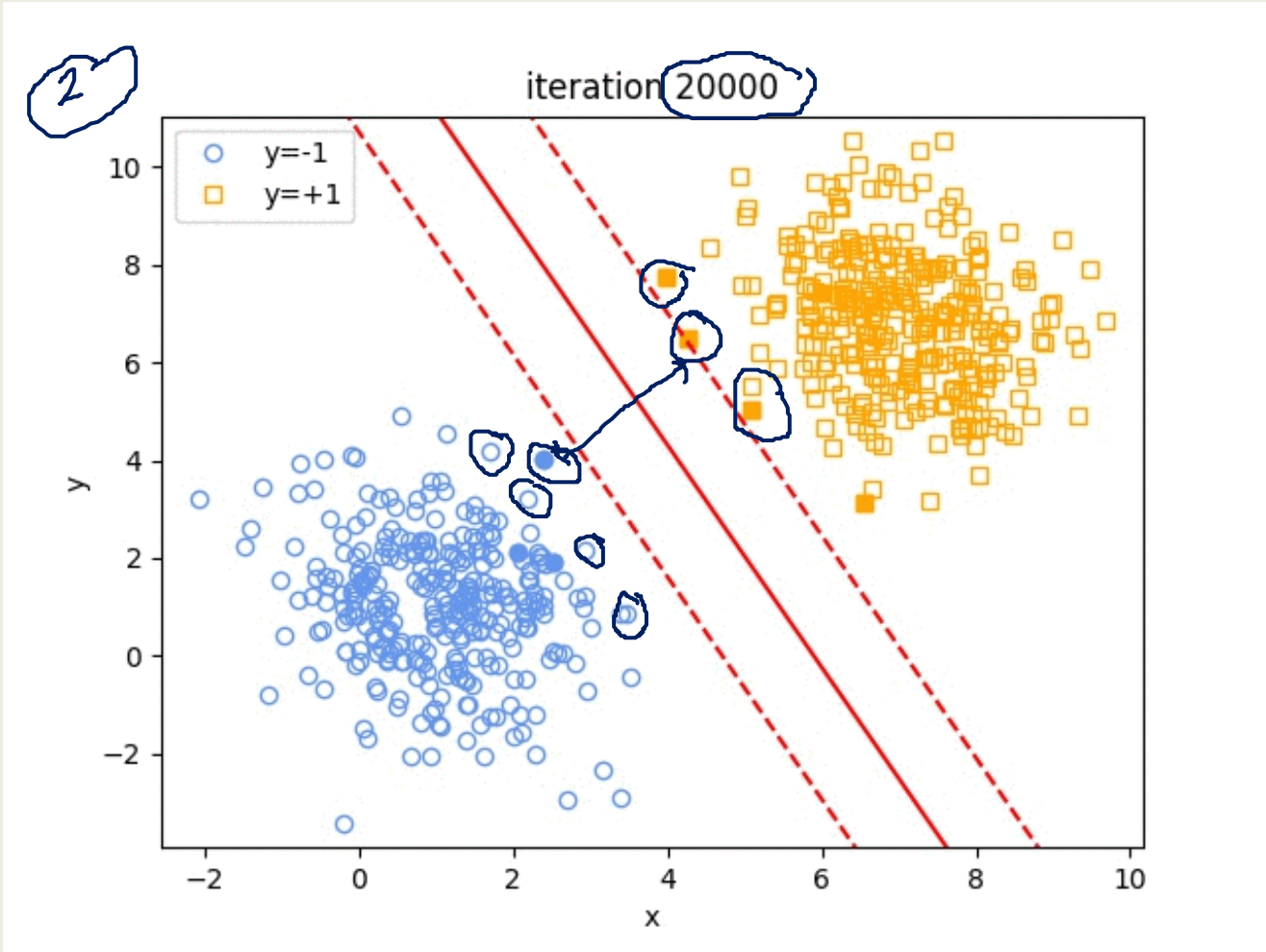
2. Hyperplane − As we can see in the above diagram, it is a decision plane or space which is divided between a set of objects having different classes.

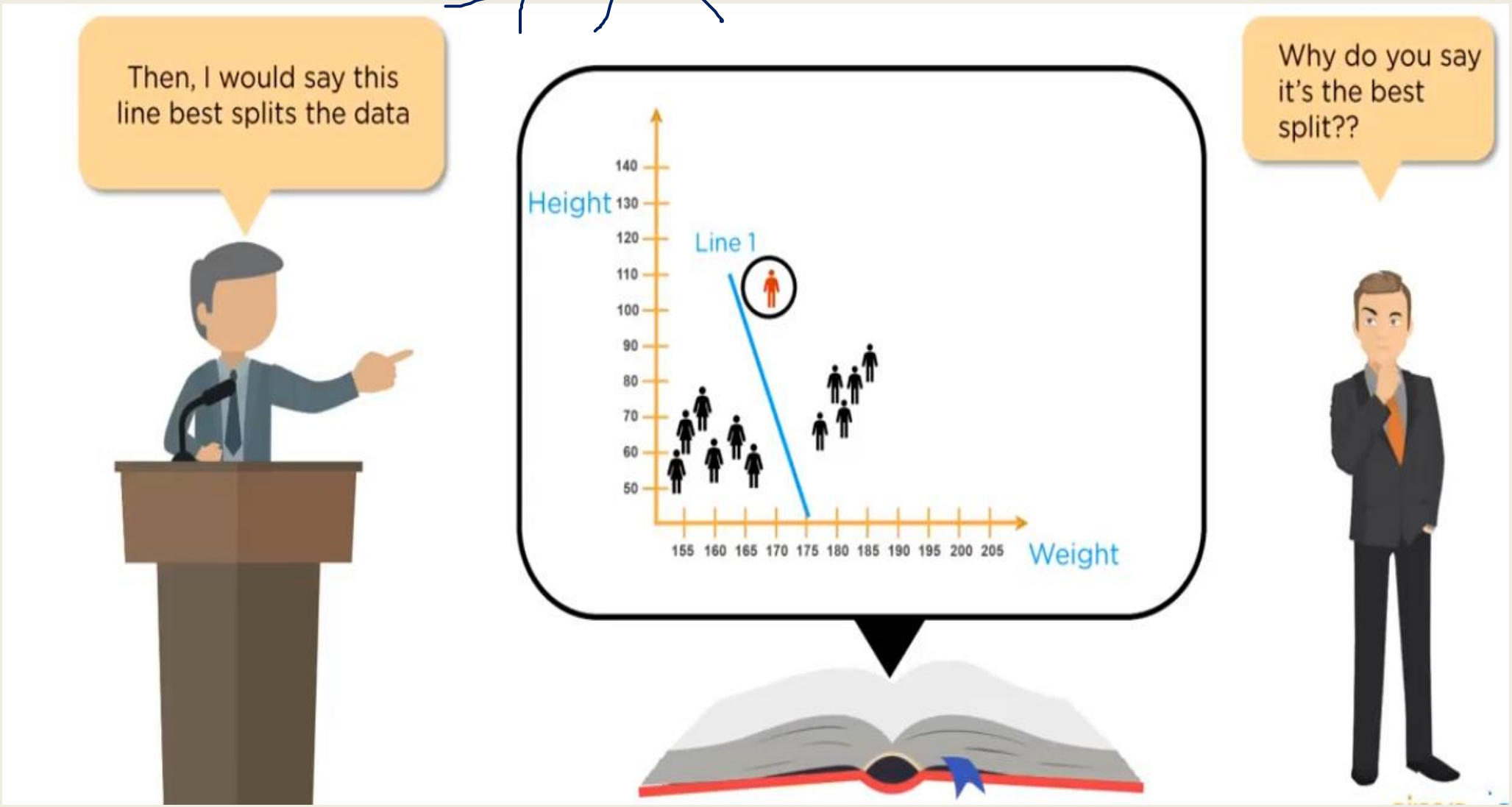
3. Margin − It may be defined as the gap between two lines on the closet data points of different classes. It can be calculated as the perpendicular distance from the line to the support vectors. Large margin is considered as a good margin and small margin is considered as a bad margin.

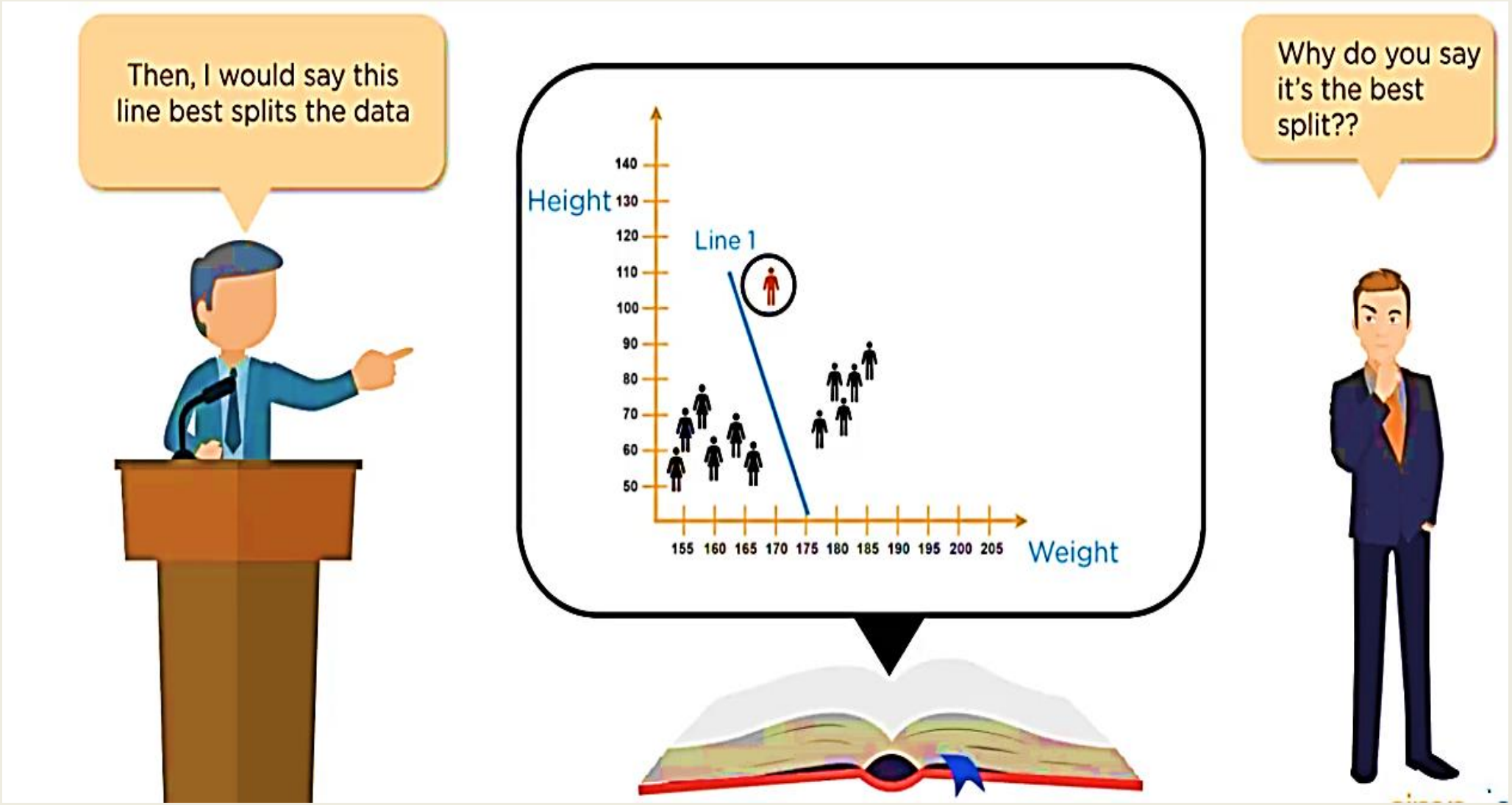

3.4, 10.9
4.2, 13.5
3.4, 10.8
5.4, 17.4
-1.8, -5.7
-2.9, -9.4
-1.7, -5.6
0.8, 2.5
2.7, 8.6
2.3, 7.3
1, 3.2
4.5, 14.5
1.1, 3.5
4.3, 14
-3.3, -10.7
1.6, 5
3, 9.7
2.3, 7.3
2.3, 7.2
-2.5, -8.1
1.9, 6
2.7, 8.5
1.9, 6
-6.7, -21.6
-6.4, -20.6
2.1, 6.9
-2.4, -7.7
-4.5, -14.4
-3.6, -11.6
-6.4, -20.6
-4.7, -15.2

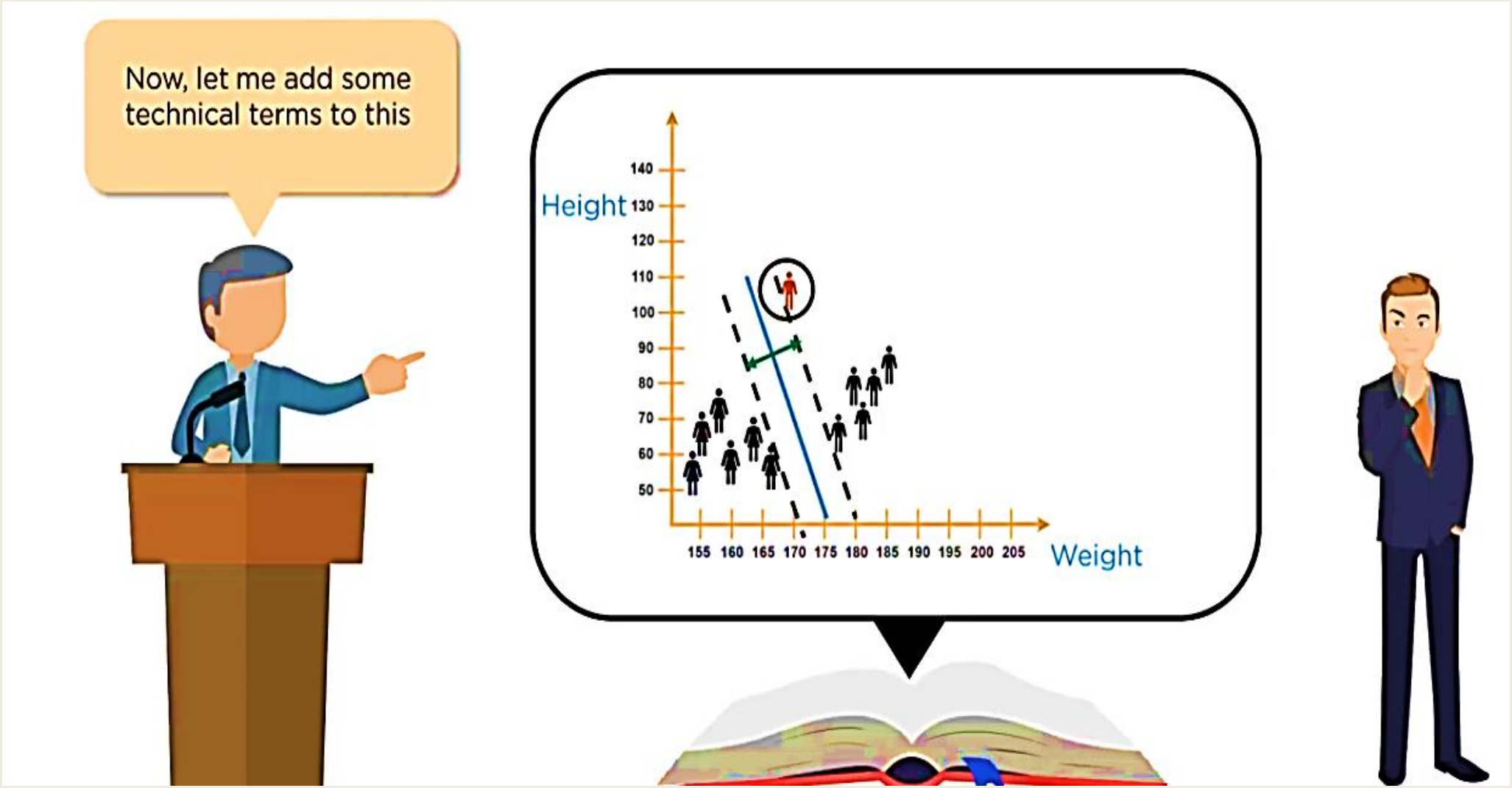

iteration 20000
y=-1
y=+1
y
x
10
8
6
4
2
0
−2
−2
0
2
4
6
8
10

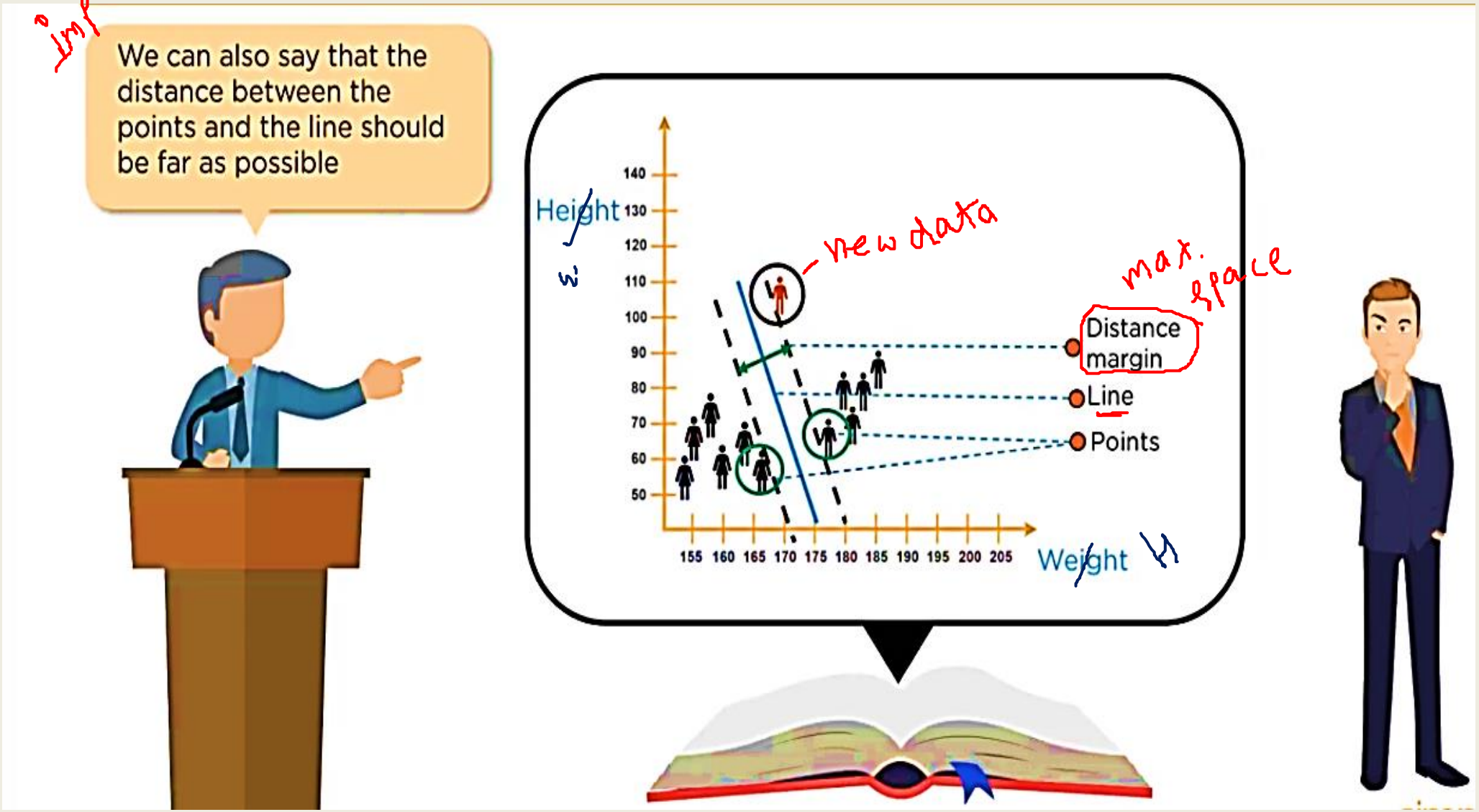

3.4, 10.9
4.2, 13.5
3.4, 10.8
5.4, 17.4
-2.9, -9.4
2.7, 8.6
2.3, 7.3
4.5, 14.5
4.3, 14
-3.3, -10.7
1.6, 5
3, 9.7
2.3, 7.3
2.3, 7.2
-2.5, -8.1
1.9, 6
2.7, 8.5
2.1, 6.6
2.1, 6.9
-6.4, -20.6
-2.4, -7.7
-4.5, -14.4
-3.6, -11.6
-6.4, -20.6
-4.7, -15.2


2
iteration 20000
y=-1
y=+1
y
x
10
8
6
4
2
0
−2
−2
0
2
4
6
8
10


Then, I would say this line best splits the data
Height
140
130
120
110
100
90
80
70
60
50
Line 1
155 160 165 170 175 180 185 190 195 200 205
Weight
Why do you say it's the best split??


Then, I would say this line best splits the data
Height
140
130
120
110
100
90
80
70
60
50
Line 1
155 160 165 170 175 180 185 190 195 200 205
Weight
Why do you say it's the best split??


Now, let me add some technical terms to this
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight


We can also say that the distance between the points and the line should be far as possible
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
new data
max. space
Distance margin
Line
Points

A model which supports "Support vector" in ML.



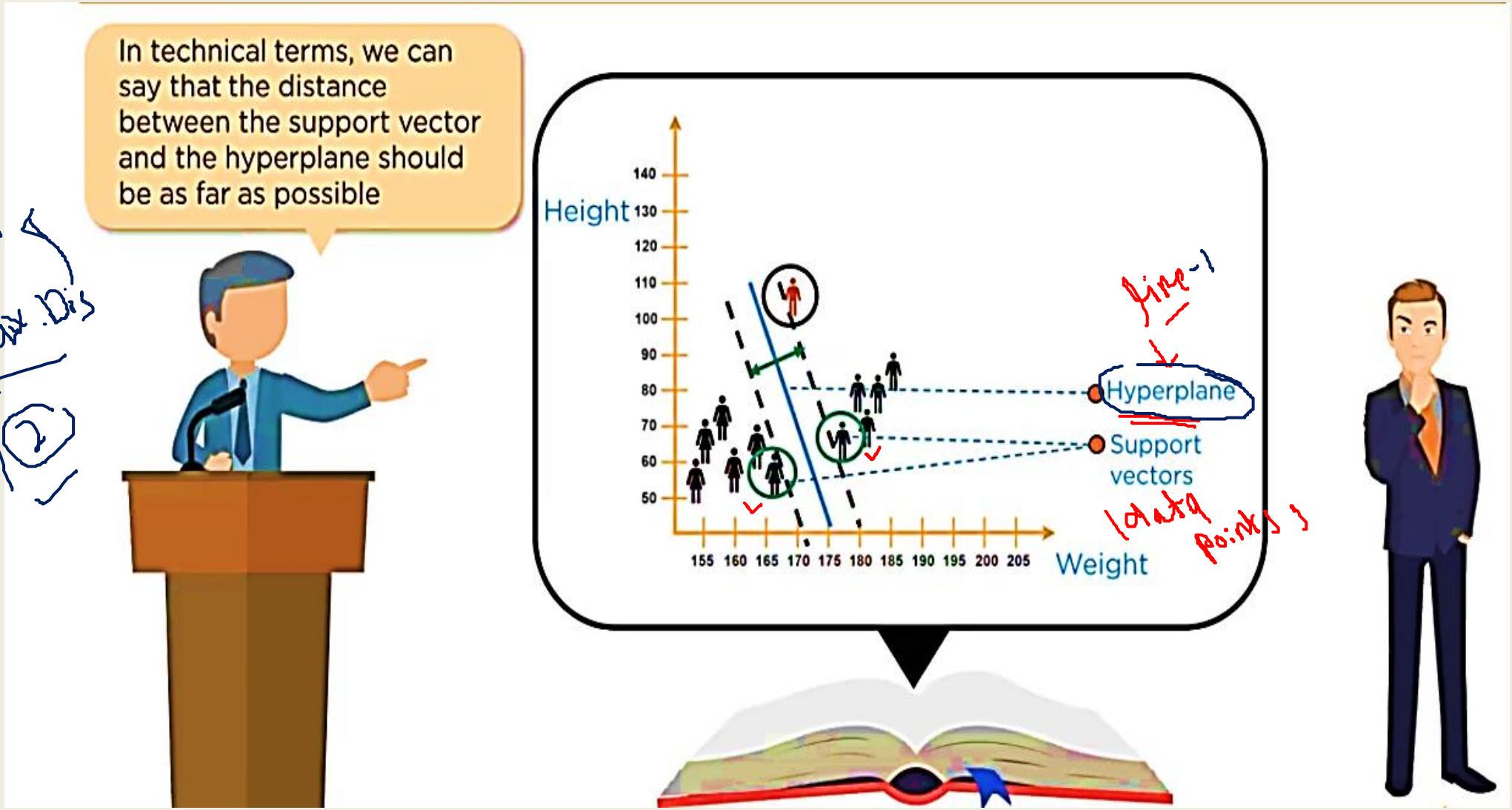
SVM

SV

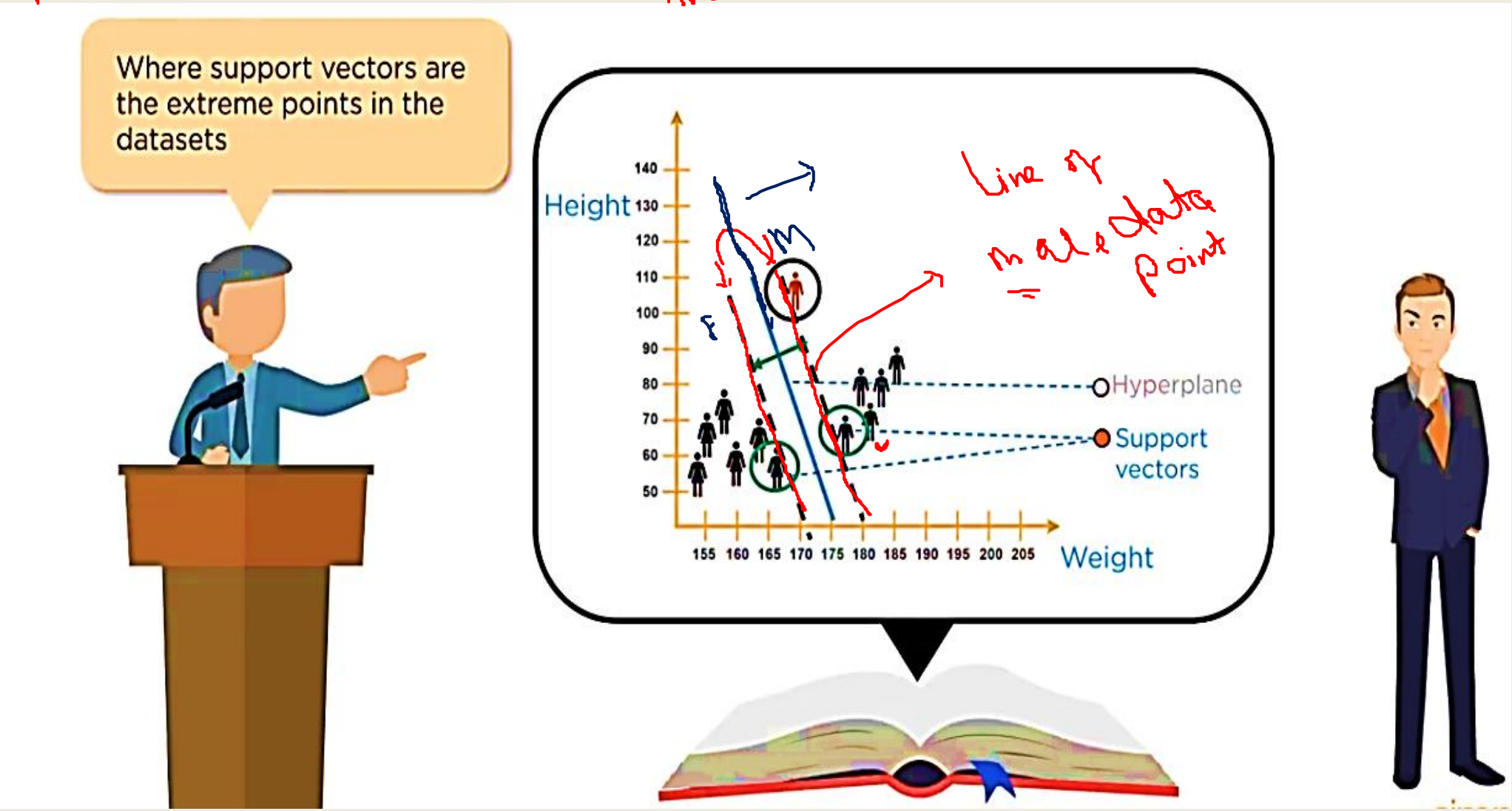
Line/HP

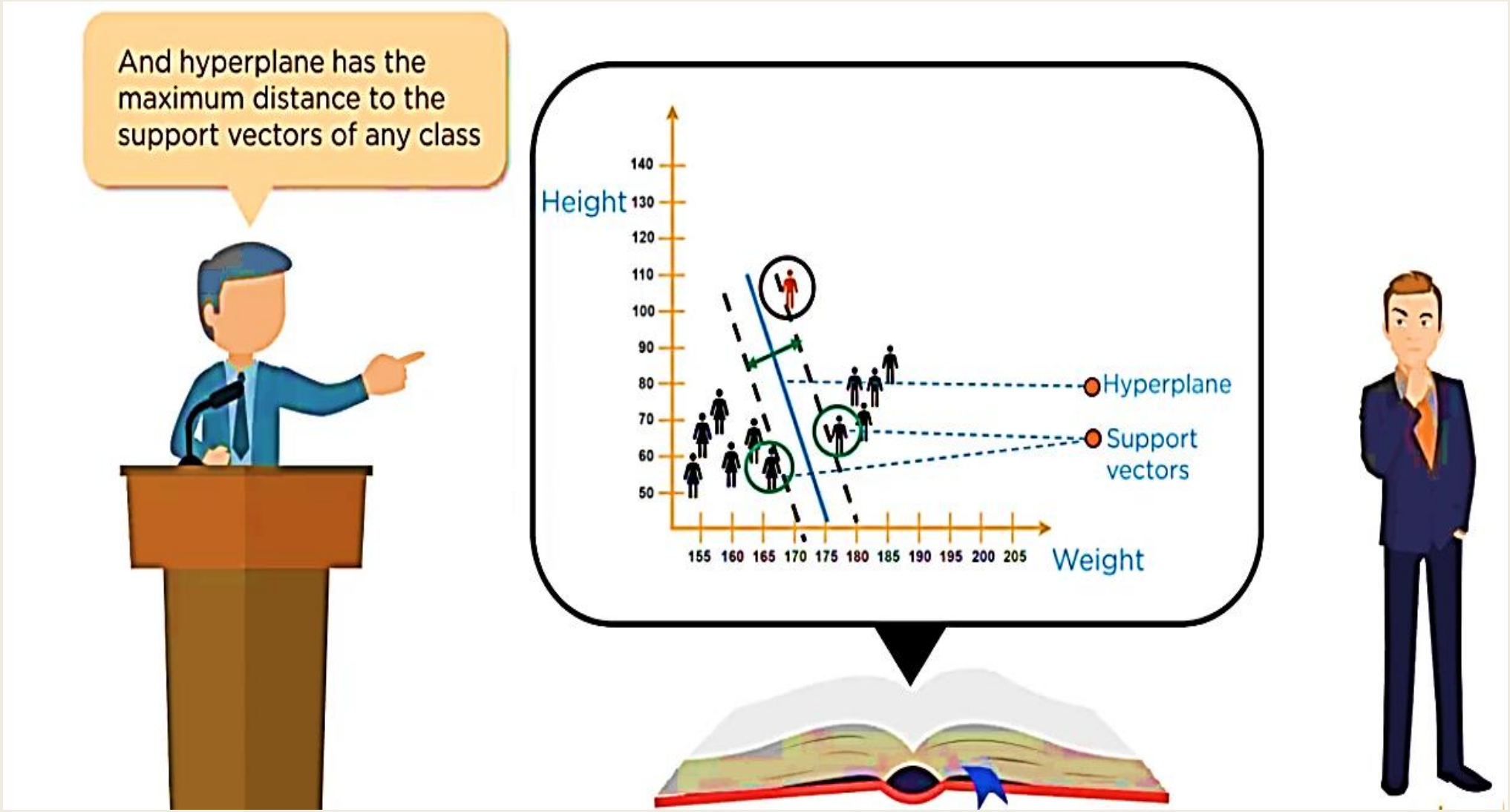
max Dis

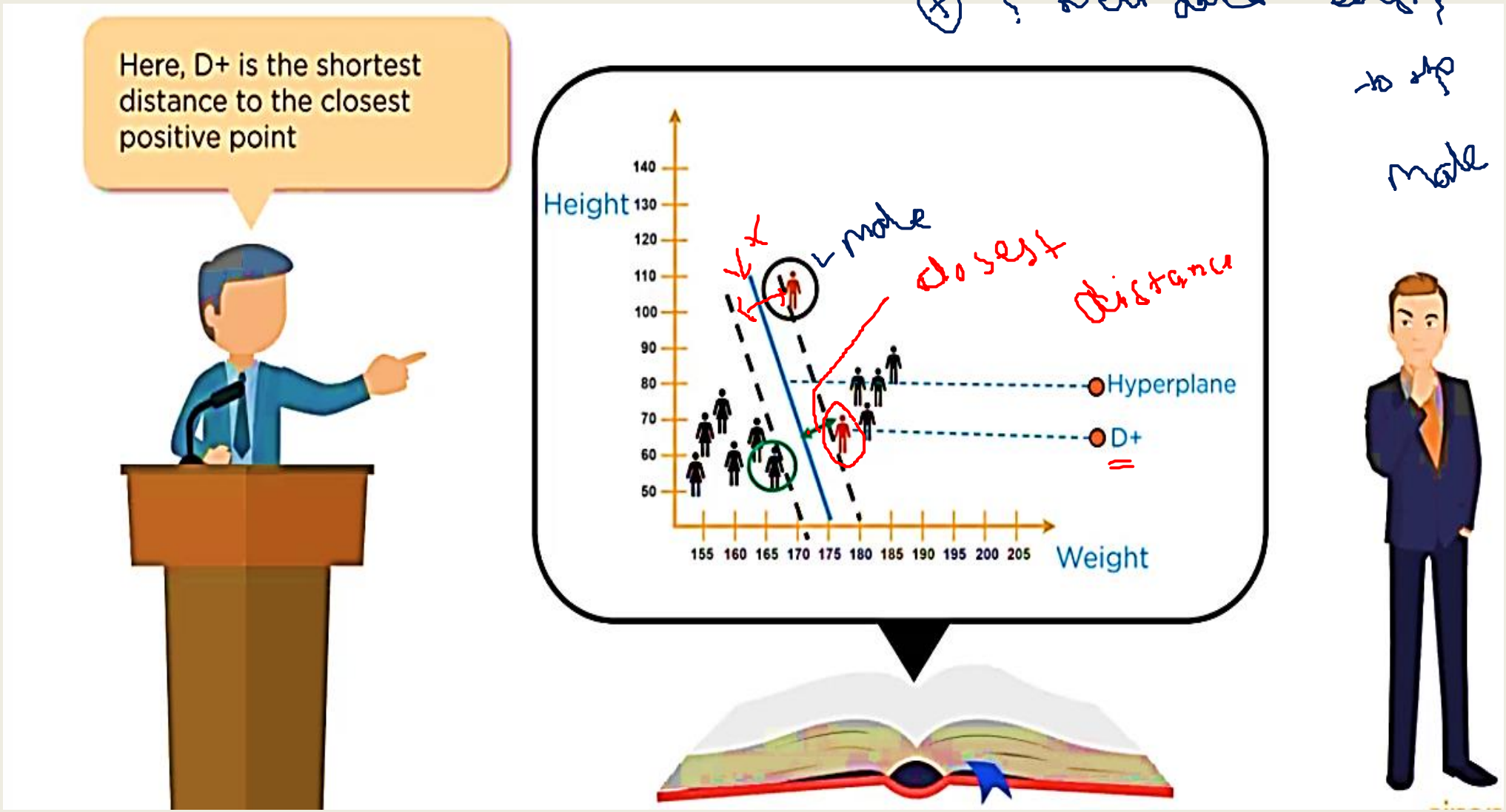
1 / 2

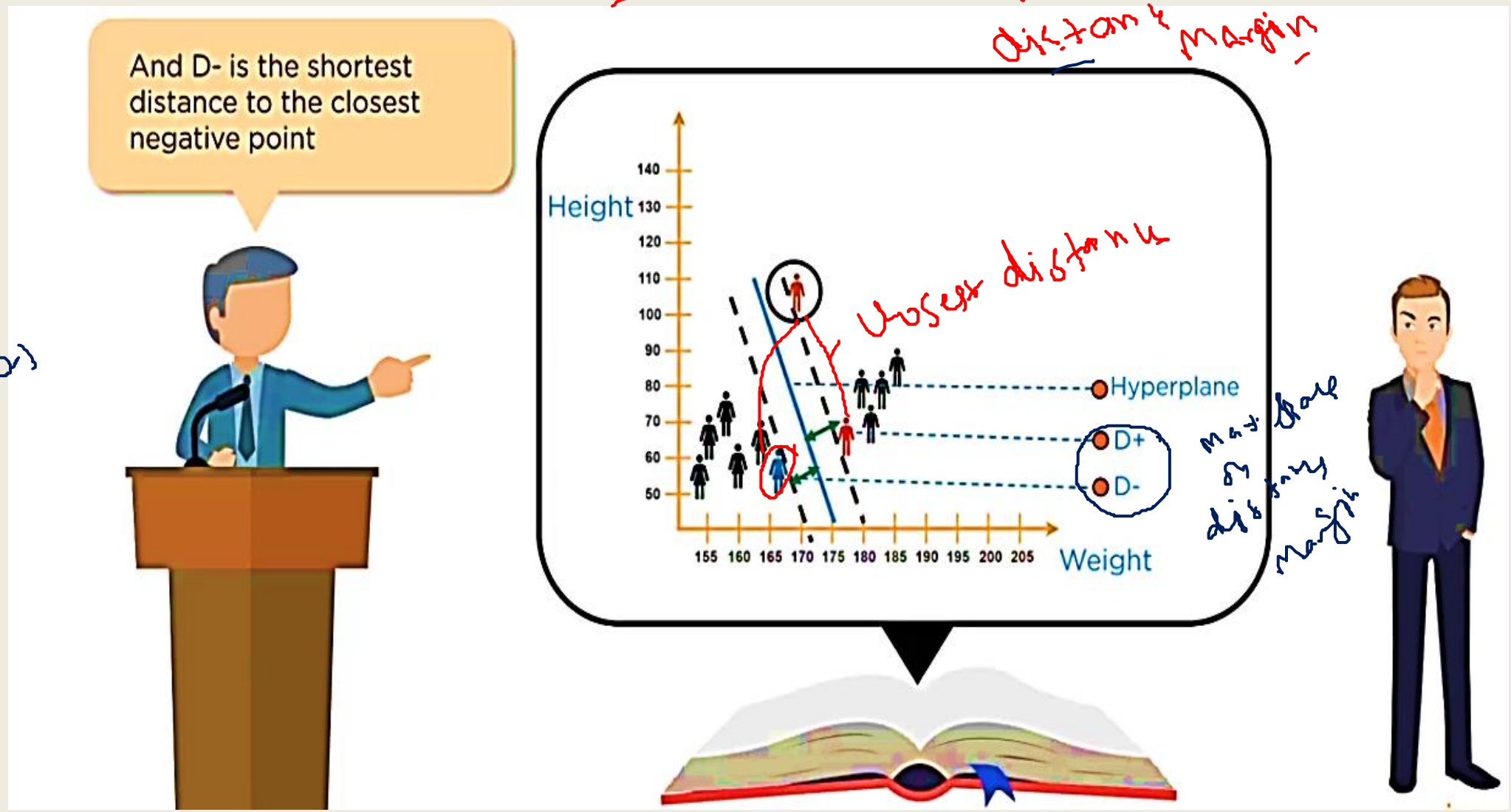


new data point is used to find of the both class. new data belong to that class.




And hyperplane has the maximum distance to the support vectors of any class
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Hyperplane
Support vectors

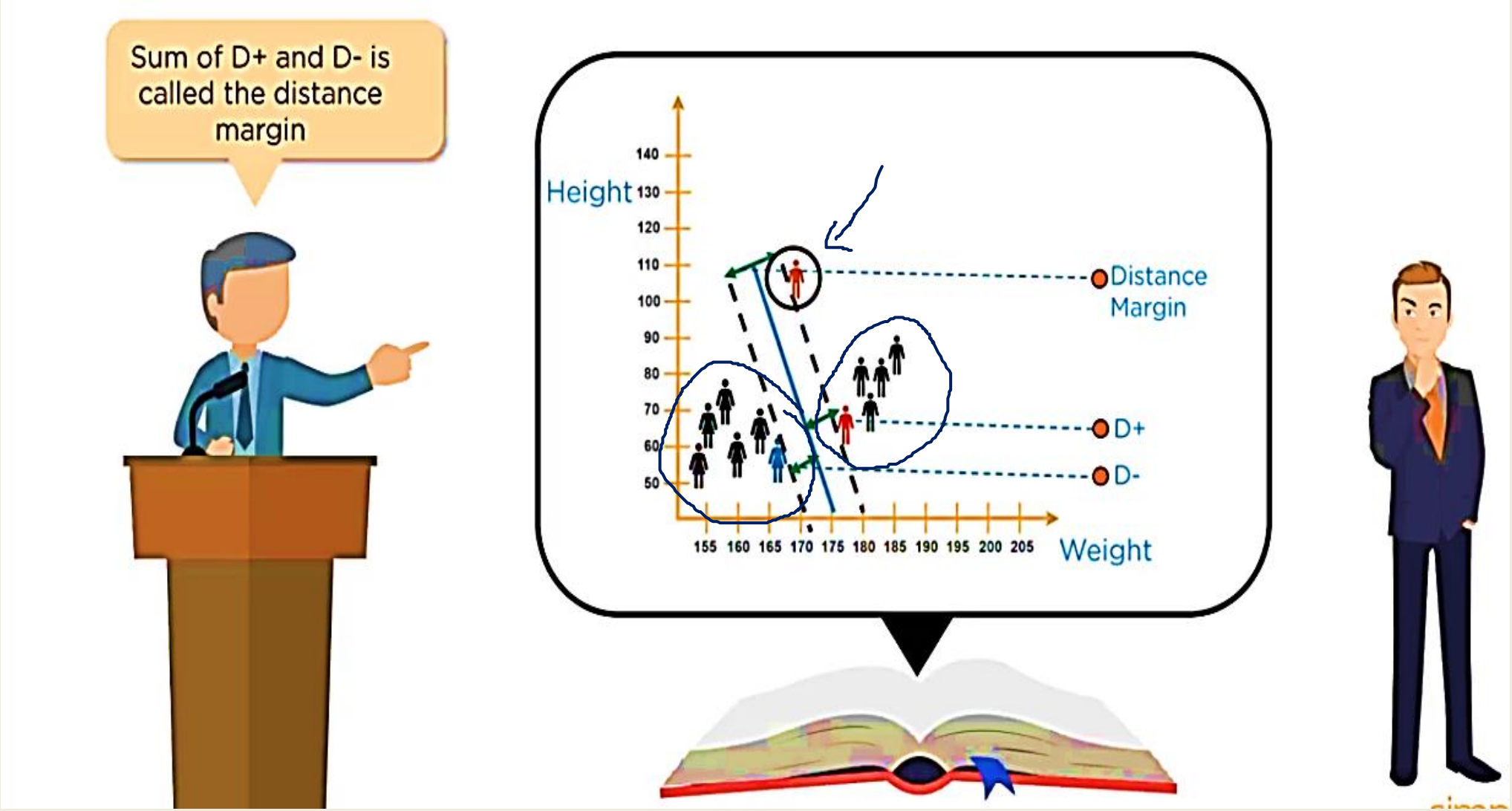

Here, D+ is the shortest distance to the closest positive point
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Hyperplane
D+

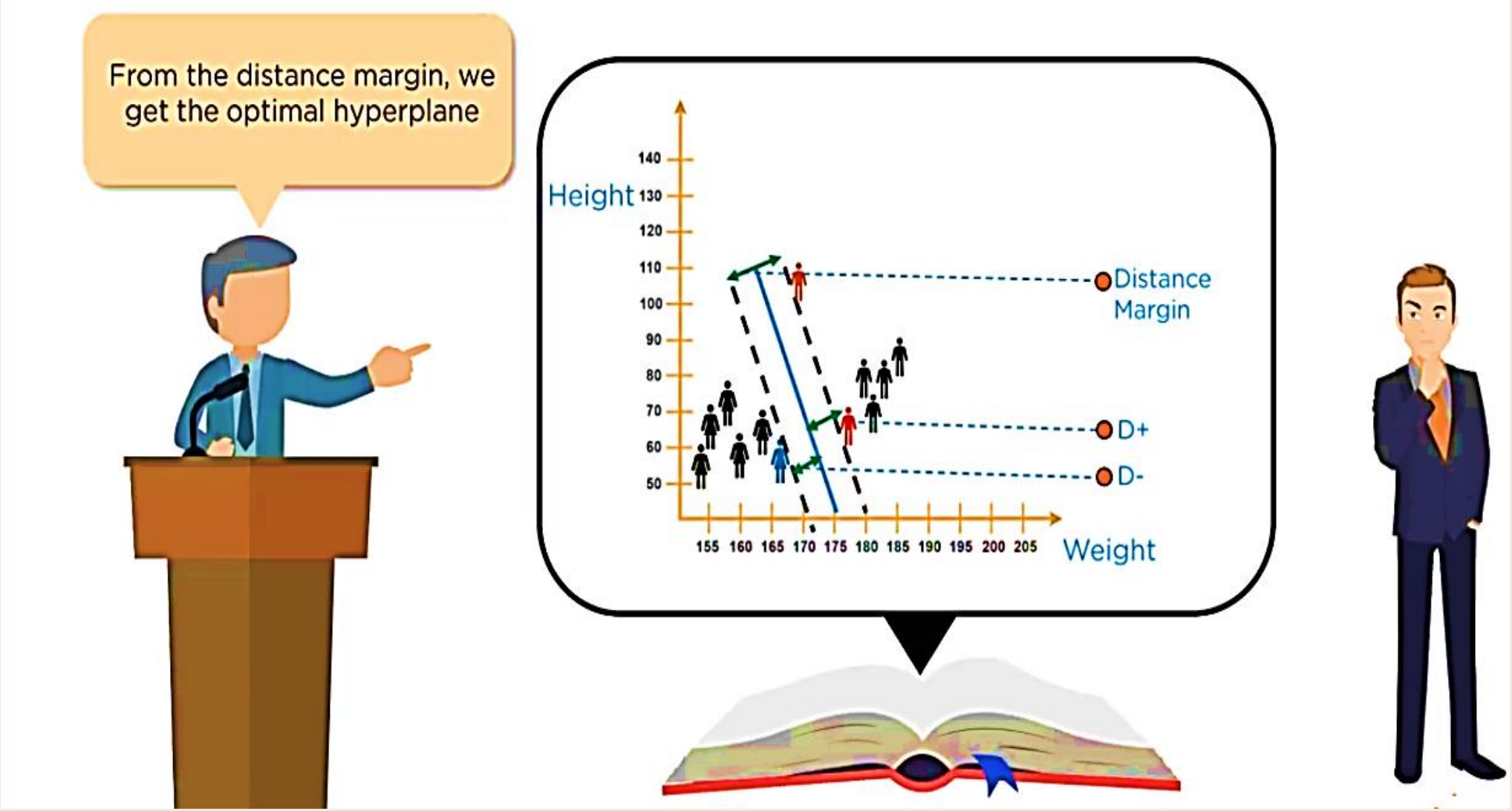

And D- is the shortest distance to the closest negative point
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Hyperplane
D+
D-

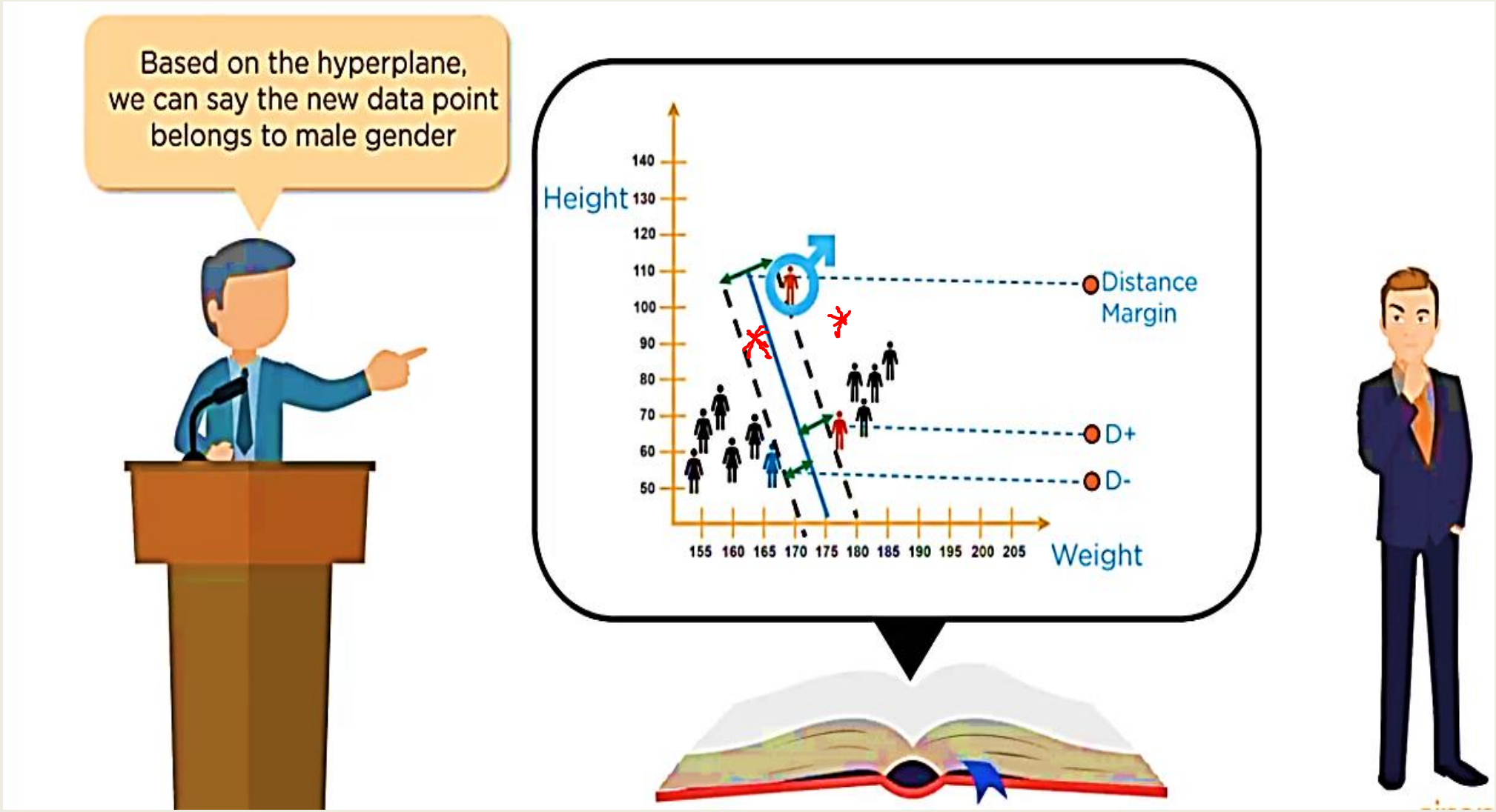
max. space / distance

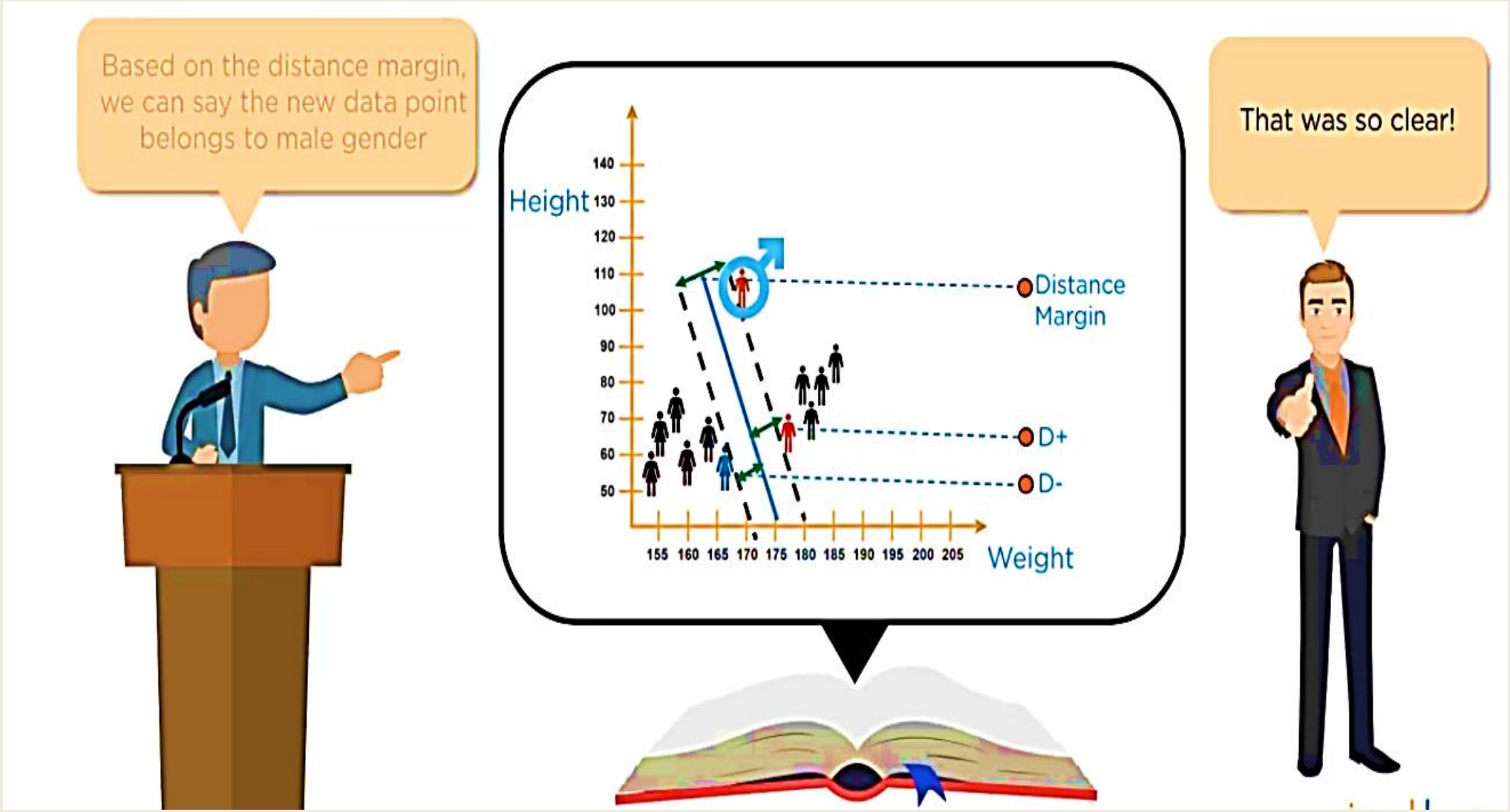
margin = (D+) + (D-)

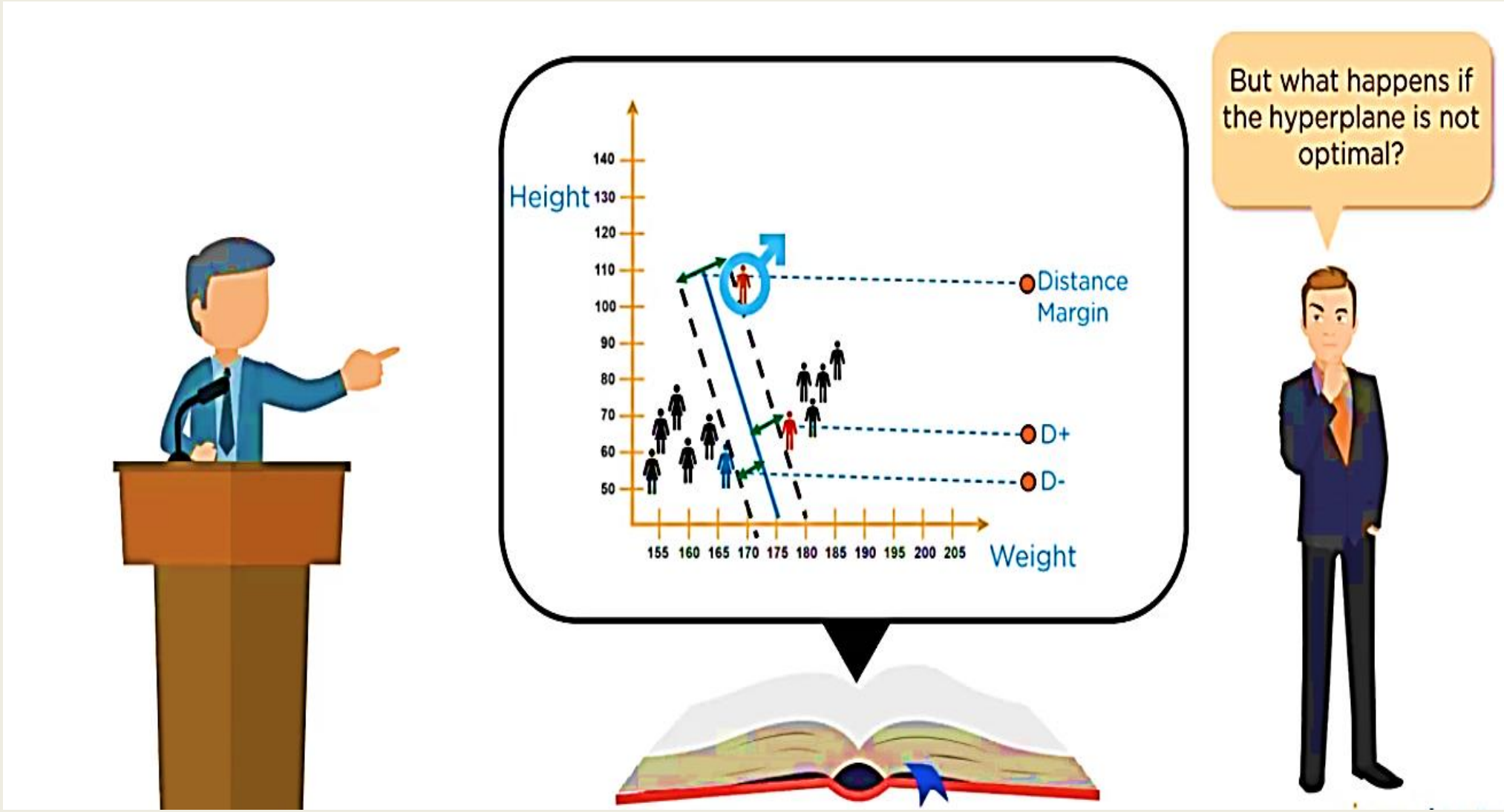


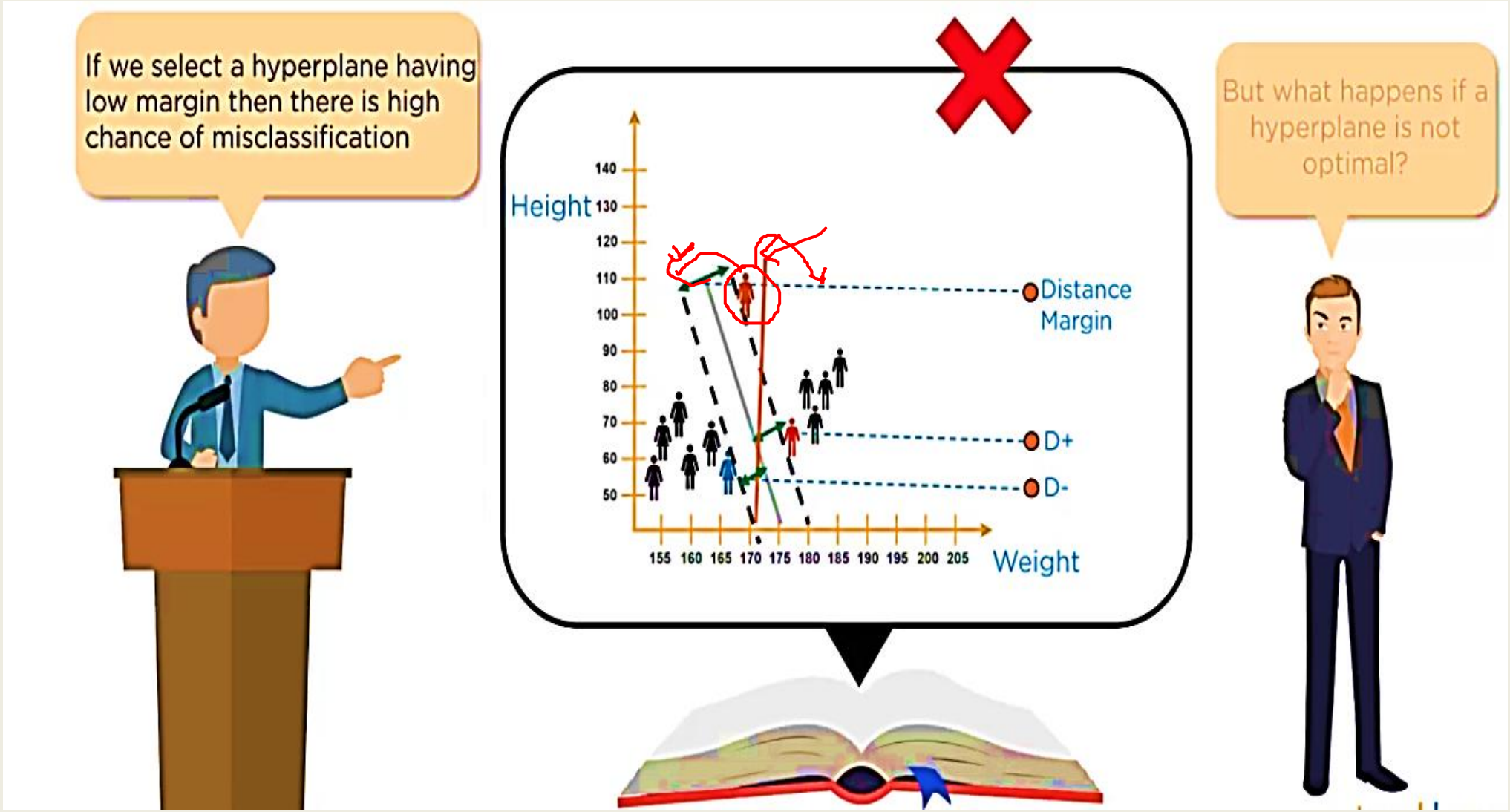


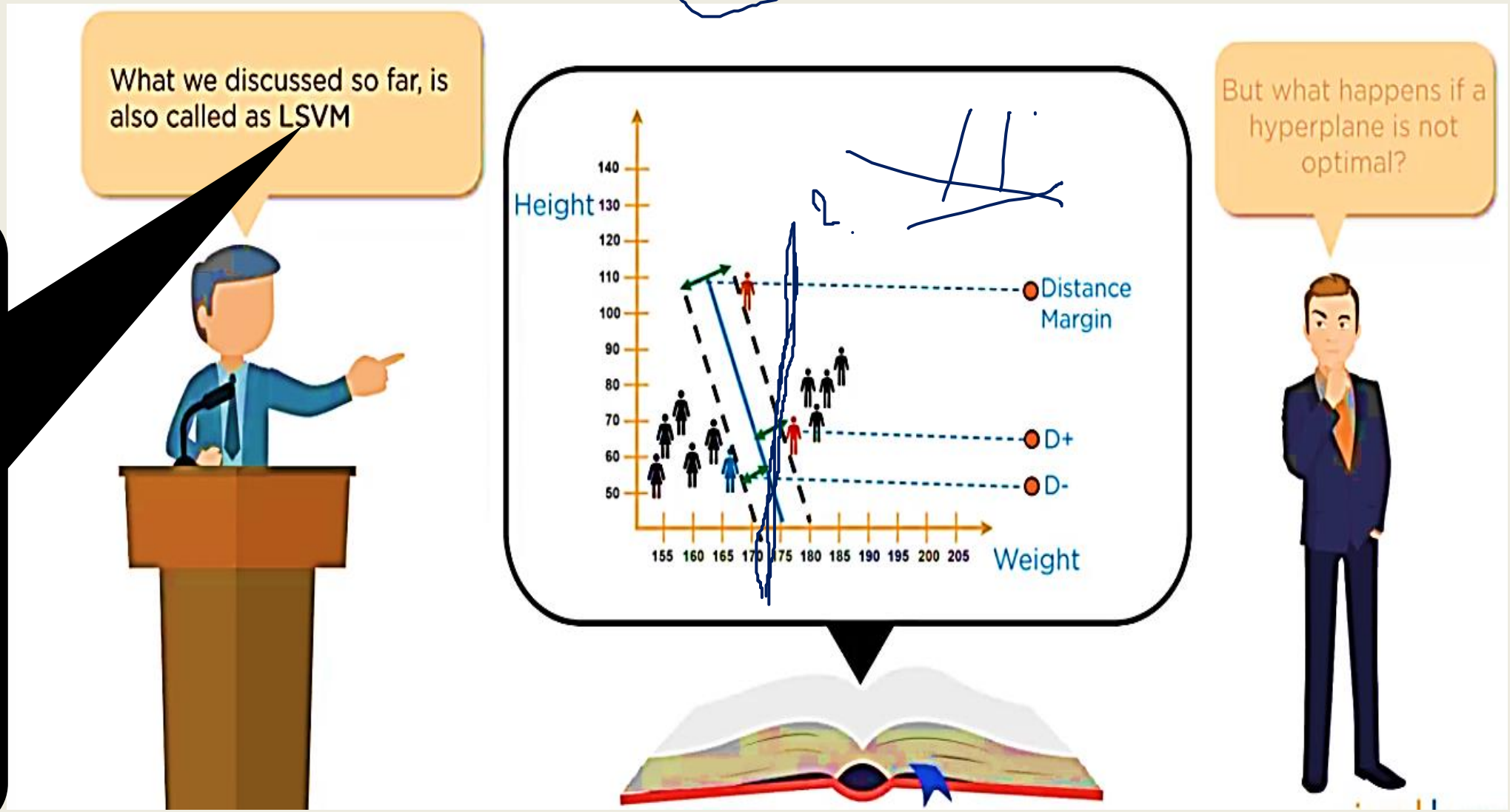

From the distance margin, we get the optimal hyperplane
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Distance Margin
D+
D-


Based on the hyperplane, we can say the new data point belongs to male gender
Height
140
130
120
110
100
90
80
70
60
50
Distance Margin
D+
D-
155 160 165 170 175 180 185 190 195 200 205
Weight


Based on the distance margin, we can say the new data point belongs to male gender
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Distance Margin
D+
D-
That was so clear!

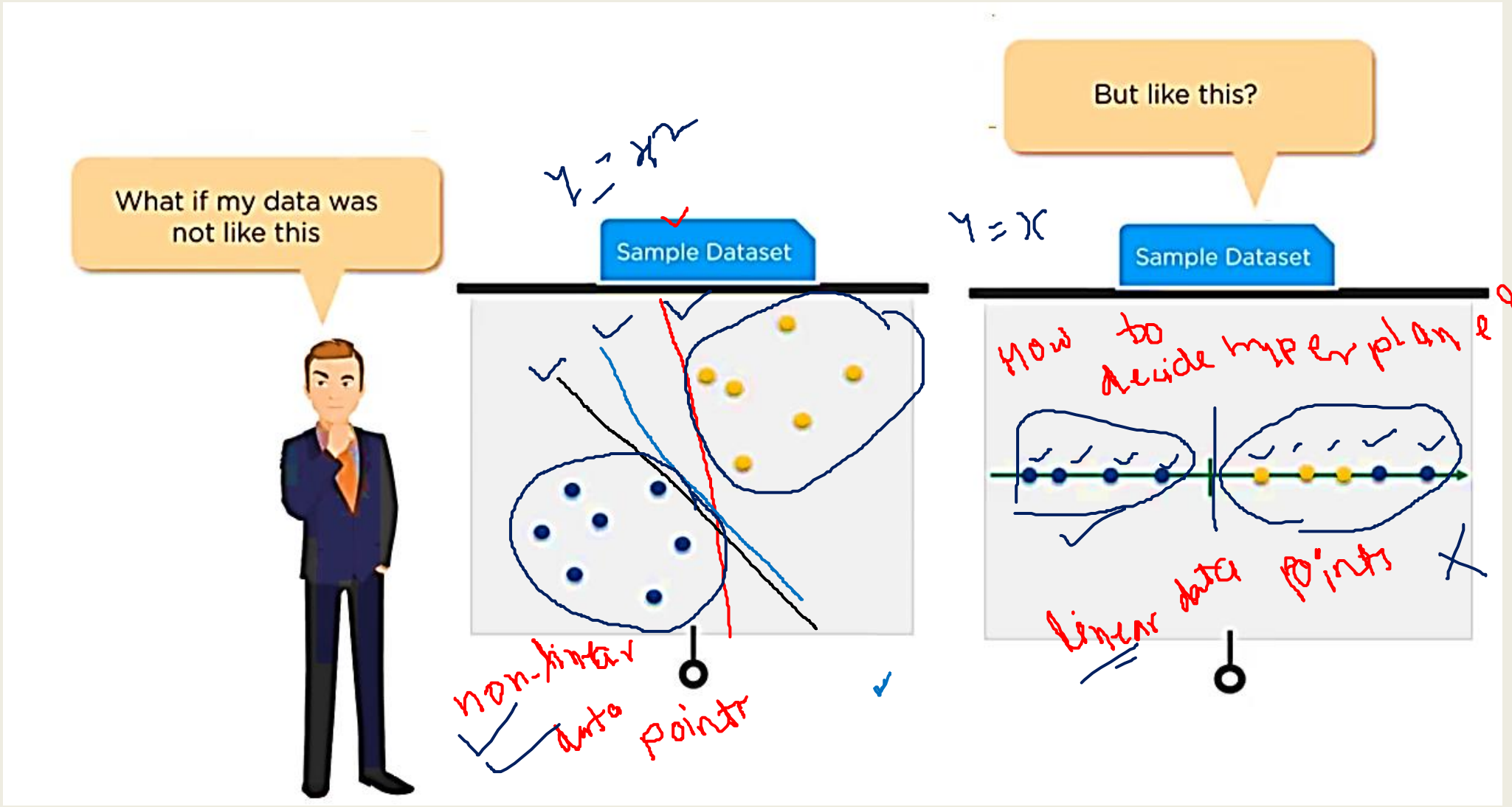

But what happens if the hyperplane is not optimal?
Height
140
130
120
110
100
90
80
70
60
50
Distance Margin
D+
D-
155 160 165 170 175 180 185 190 195 200 205
Weight

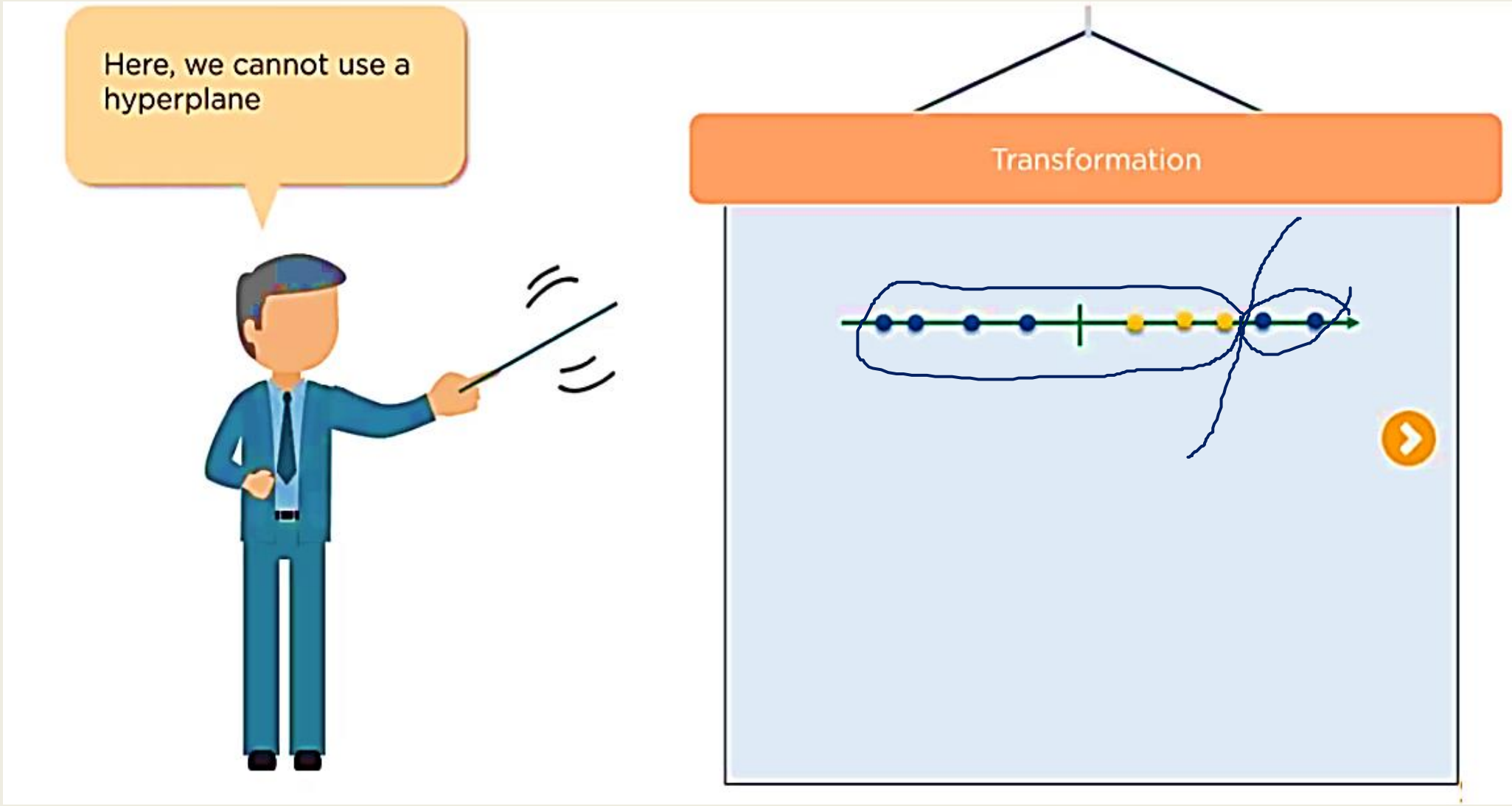

If we select a hyperplane having low margin then there is high chance of misclassification
Height
140
130
120
110
100
90
80
70
60
50
155 160 165 170 175 180 185 190 195 200 205
Weight
Distance Margin
D+
D-
But what happens if a hyperplane is not optimal?

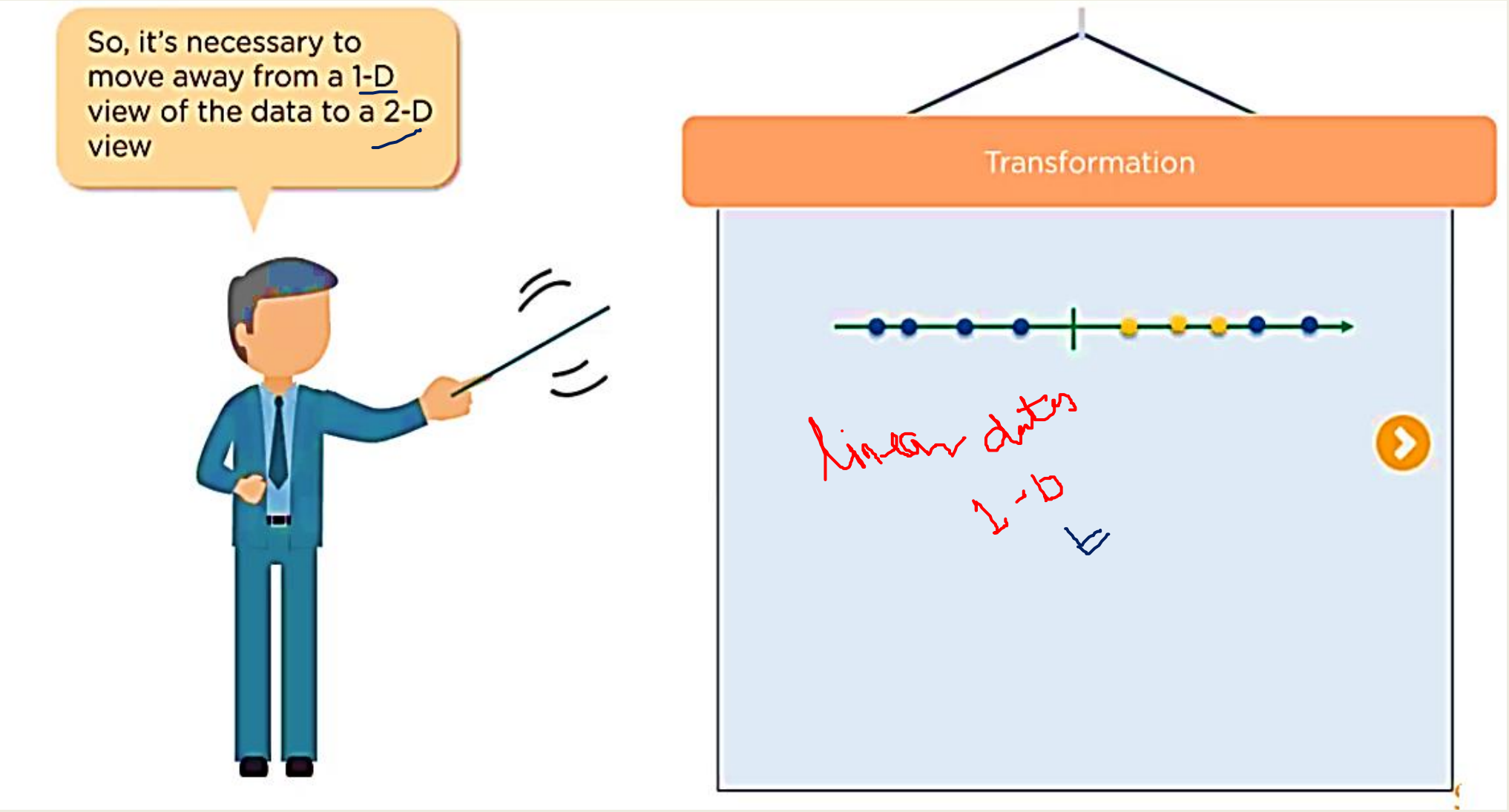



LSVM Lagrangian Support Vector Machine) ( is a fast technique for training support vector machines (SVMs), based on a simple iterative approach.

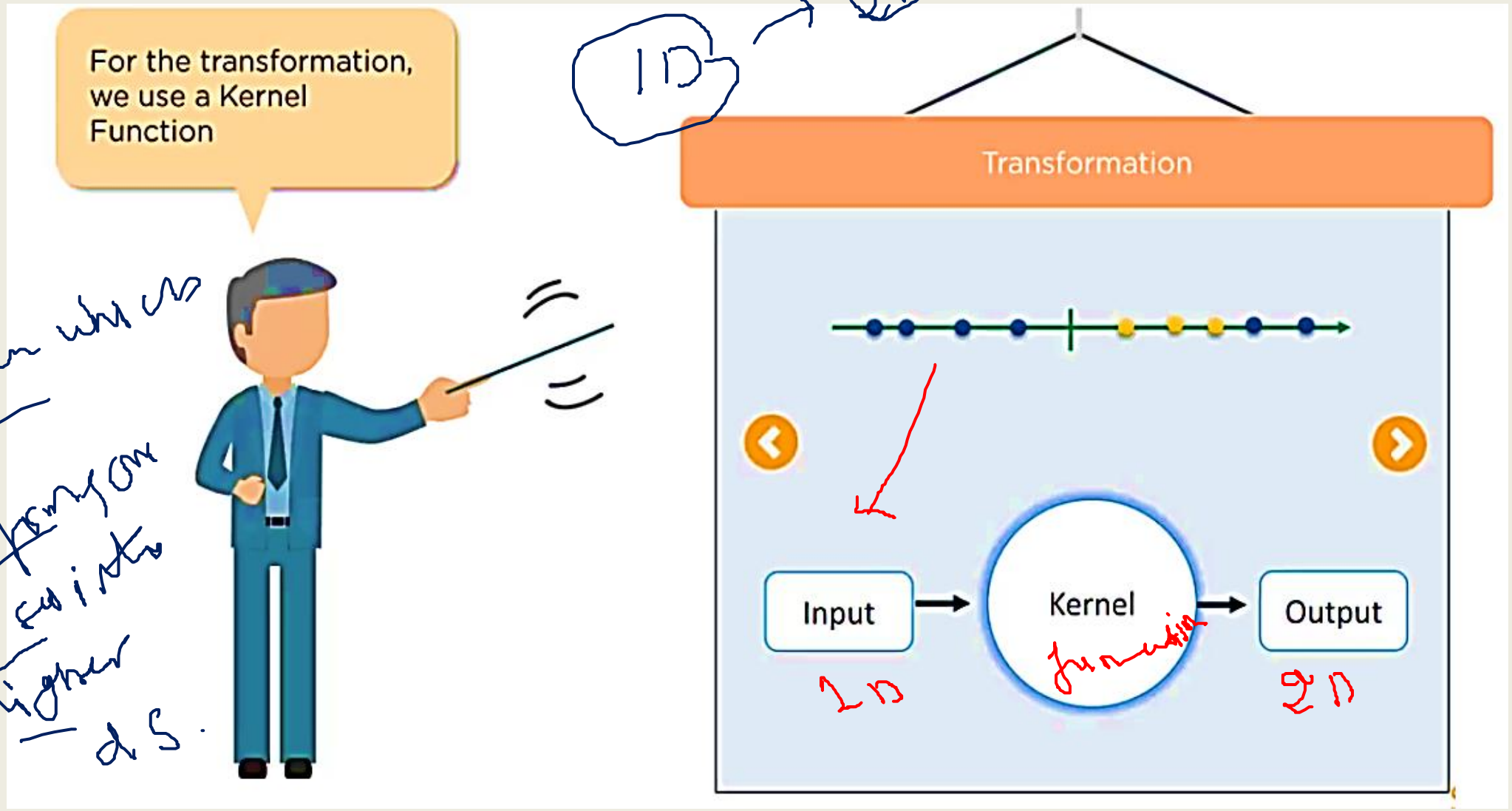
# Understanding Support Vector Machine


What if my data was not like this
Sample Dataset
But like this?
Sample Dataset
How to decide hyperplane ?
Linear data points
non-linear data points

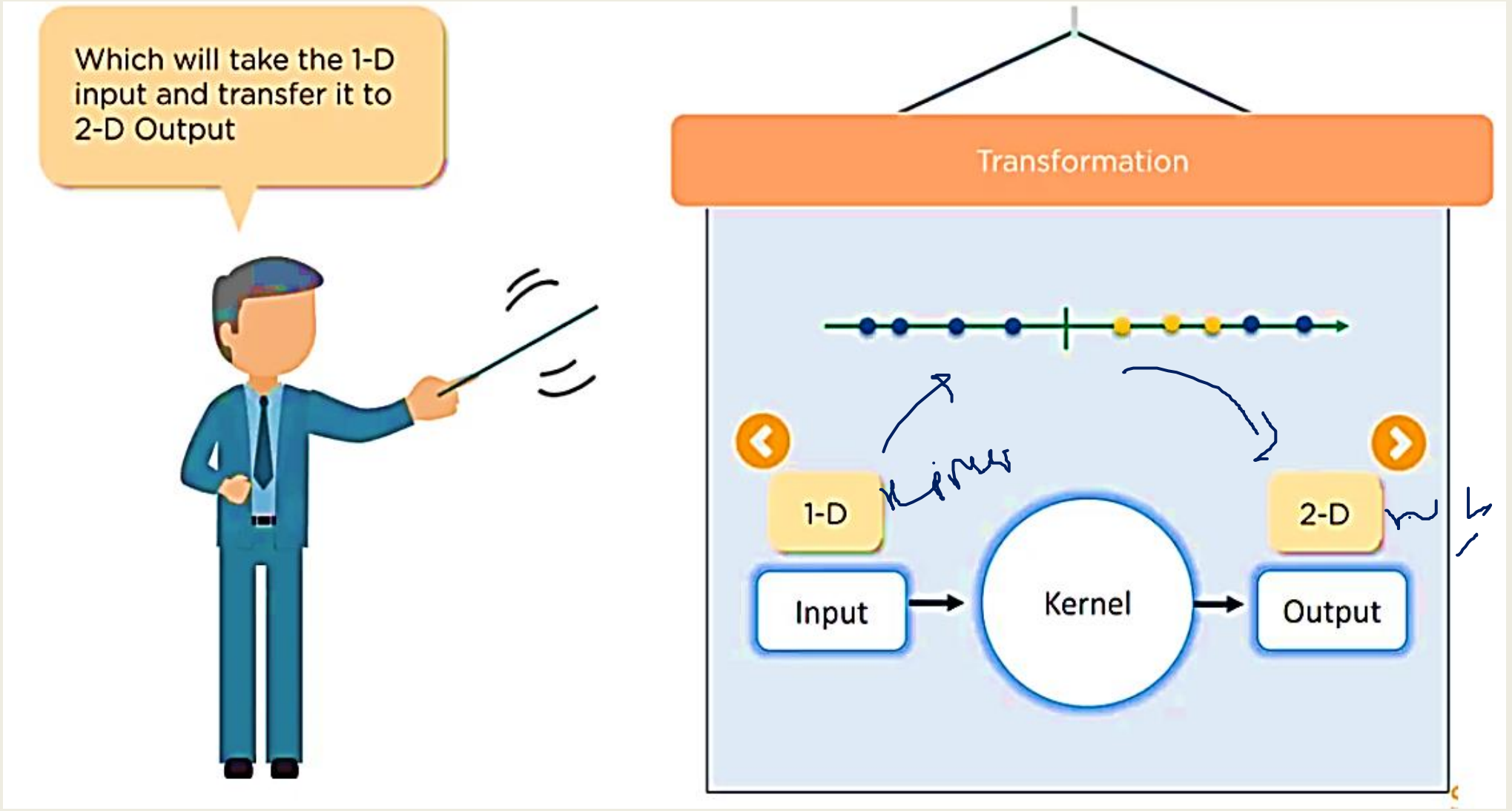

Here, we cannot use a hyperplane
Transformation

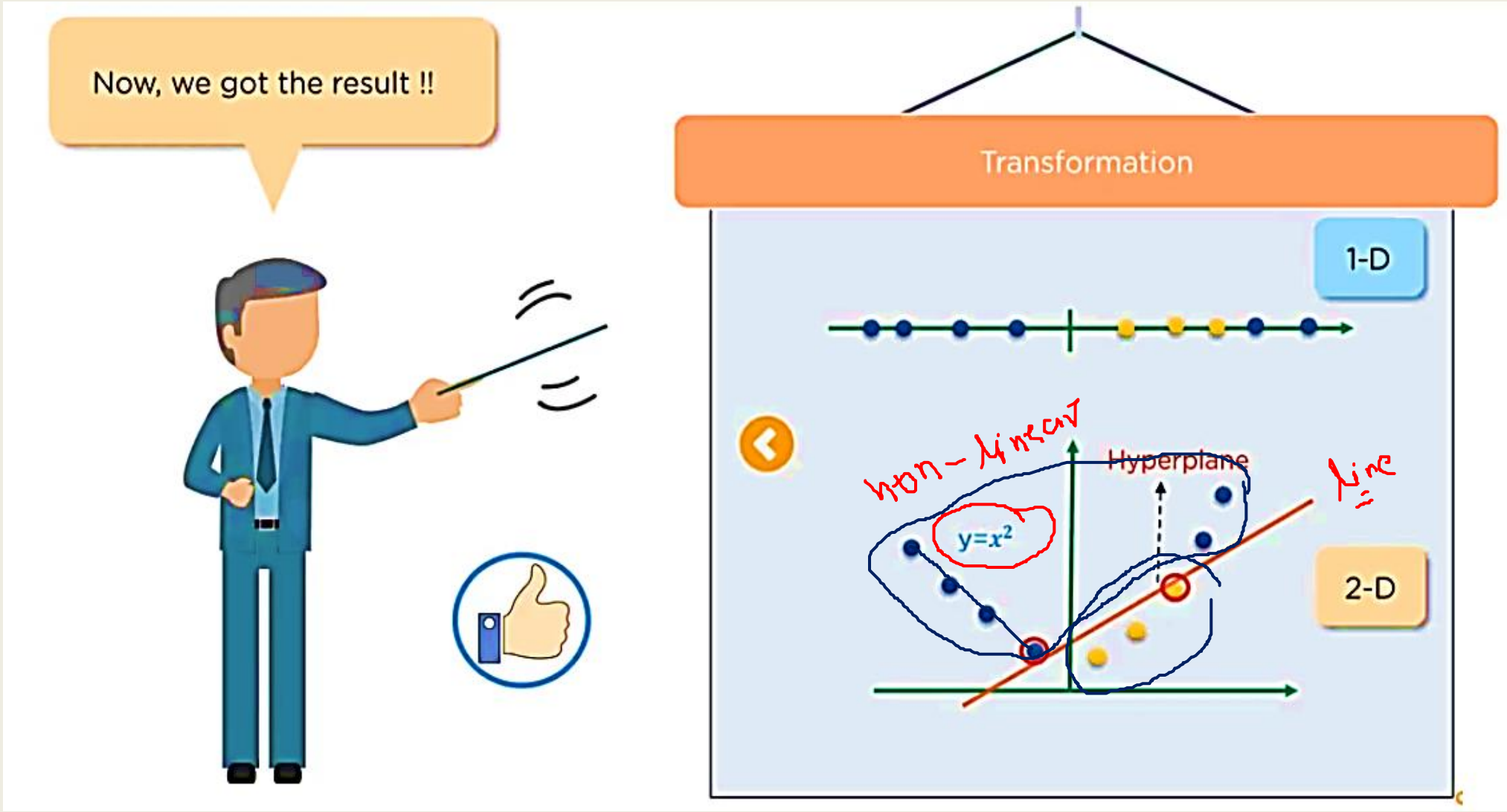

So, it's necessary to move away from a 1-D view of the data to a 2-D view
Transformation
linear data
1-D

1D - Dataset to
2D - Dataset

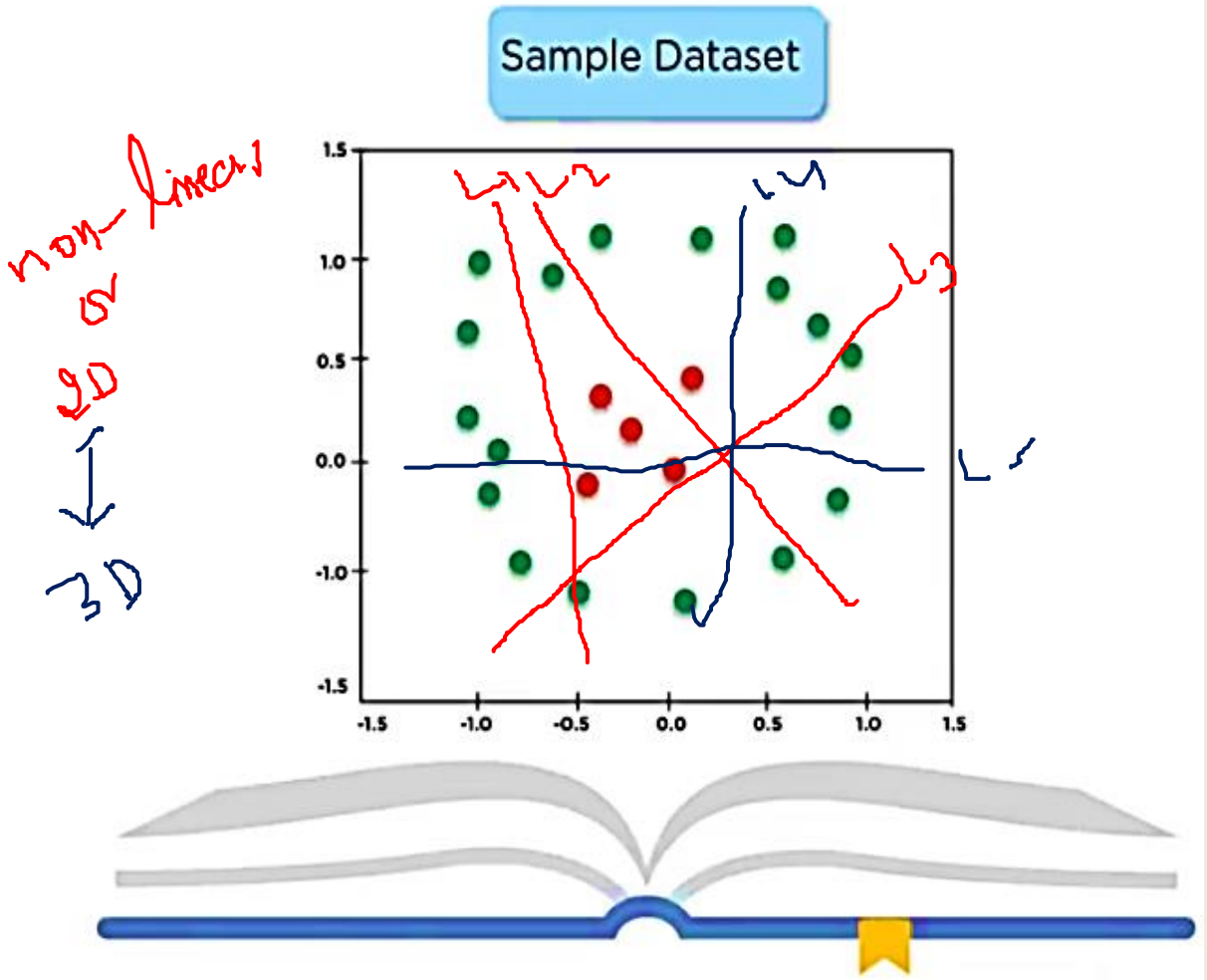


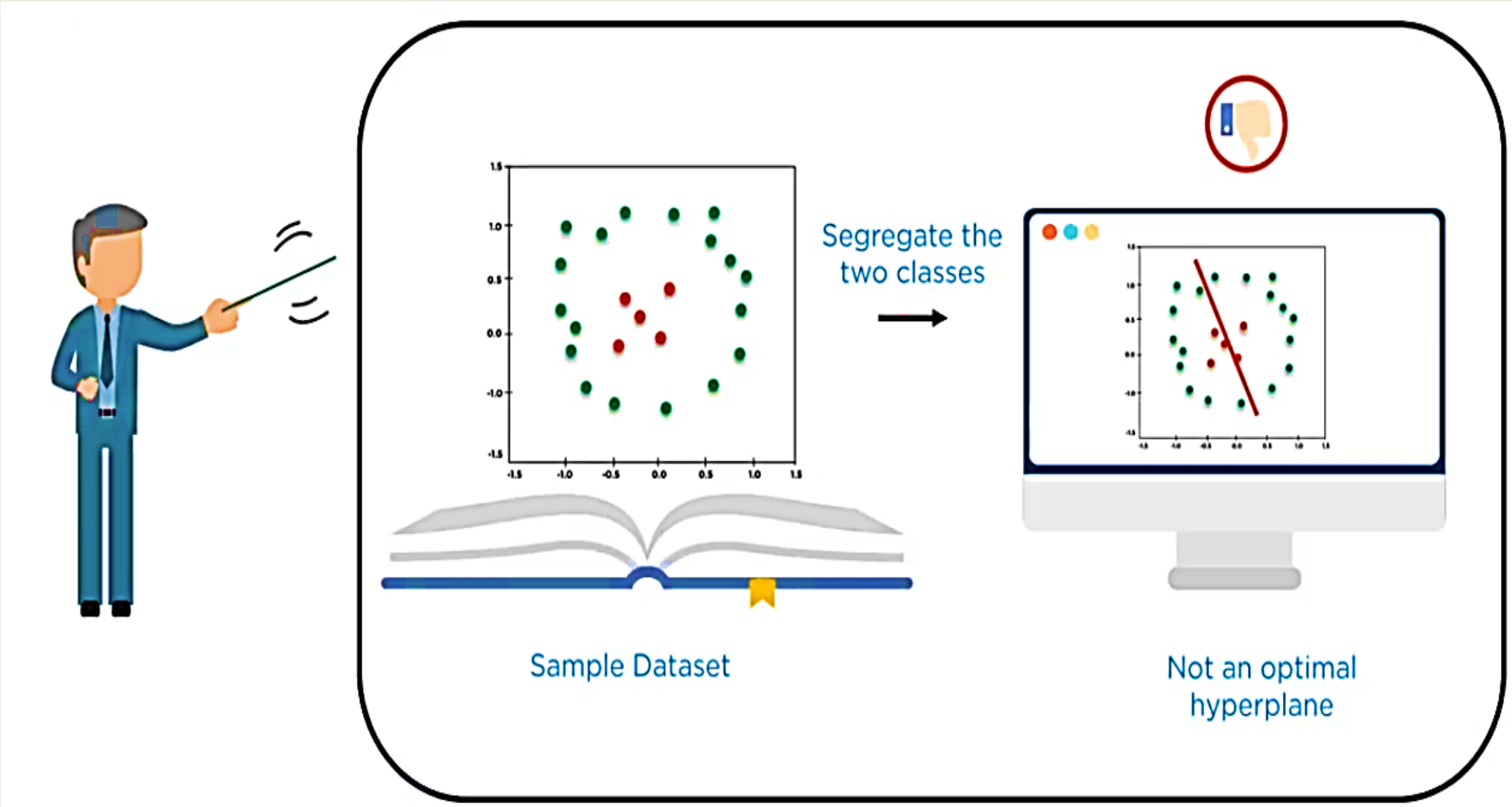

A function which transforms data existing in higher d.s.

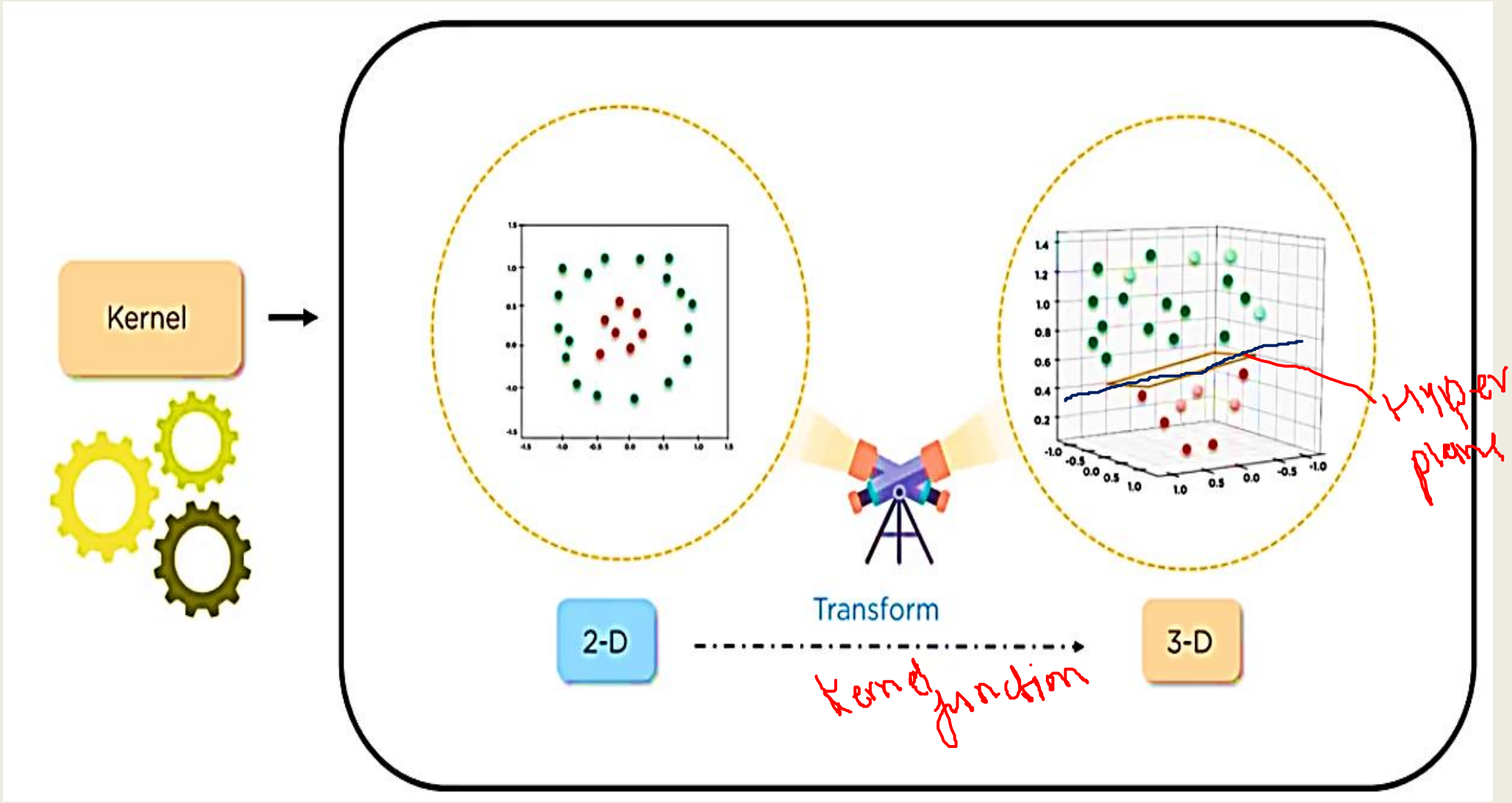

Which will take the 1-D input and transfer it to 2-D Output
Transformation
1-D
2-D
Input
Kernel
Output


Now, we got the result !!
Transformation
1-D
Hyperplane
y=x²
2-D


How to perform SVM
for this type of dataset?


Sample Dataset
non-linear
or
2D
3D
L1
L2
L3
L4
L5
1.5
1.0
0.5
0.0
-1.0
-1.5
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5


Segregate the two classes
Sample Dataset
Not an optimal hyperplane


Kernel
2-D
Transform
3-D
Kernel function
Hyper plane

# Points to be discussed

1. Why Support Vector Machine?
2. What is Support Vector Machine?
3. Understanding Support Vector Machine


VIT®

BHOPAL

www.vitbhopal.ac.in