

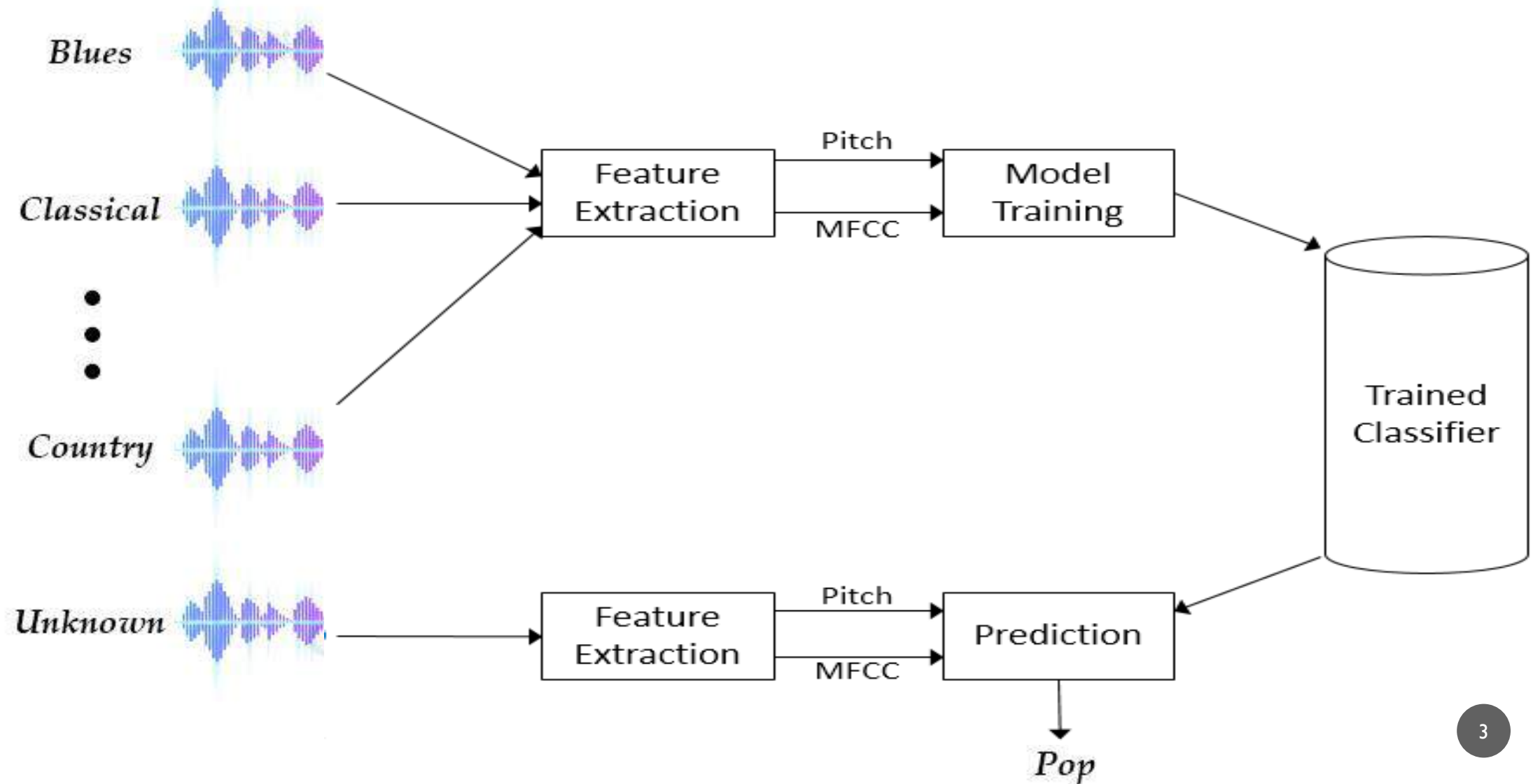
# MUSIC GENRE CLASSIFICATION

By-Abhishek Srivastav

# OUTLINE

- Overview & Objective
- Contribution
- Dataset
- Feature Extraction
- Classifier Methods
- Comparision

# OVERVIEW & OBJECTIVE



## DATASET

The GTZAN genre collection dataset:

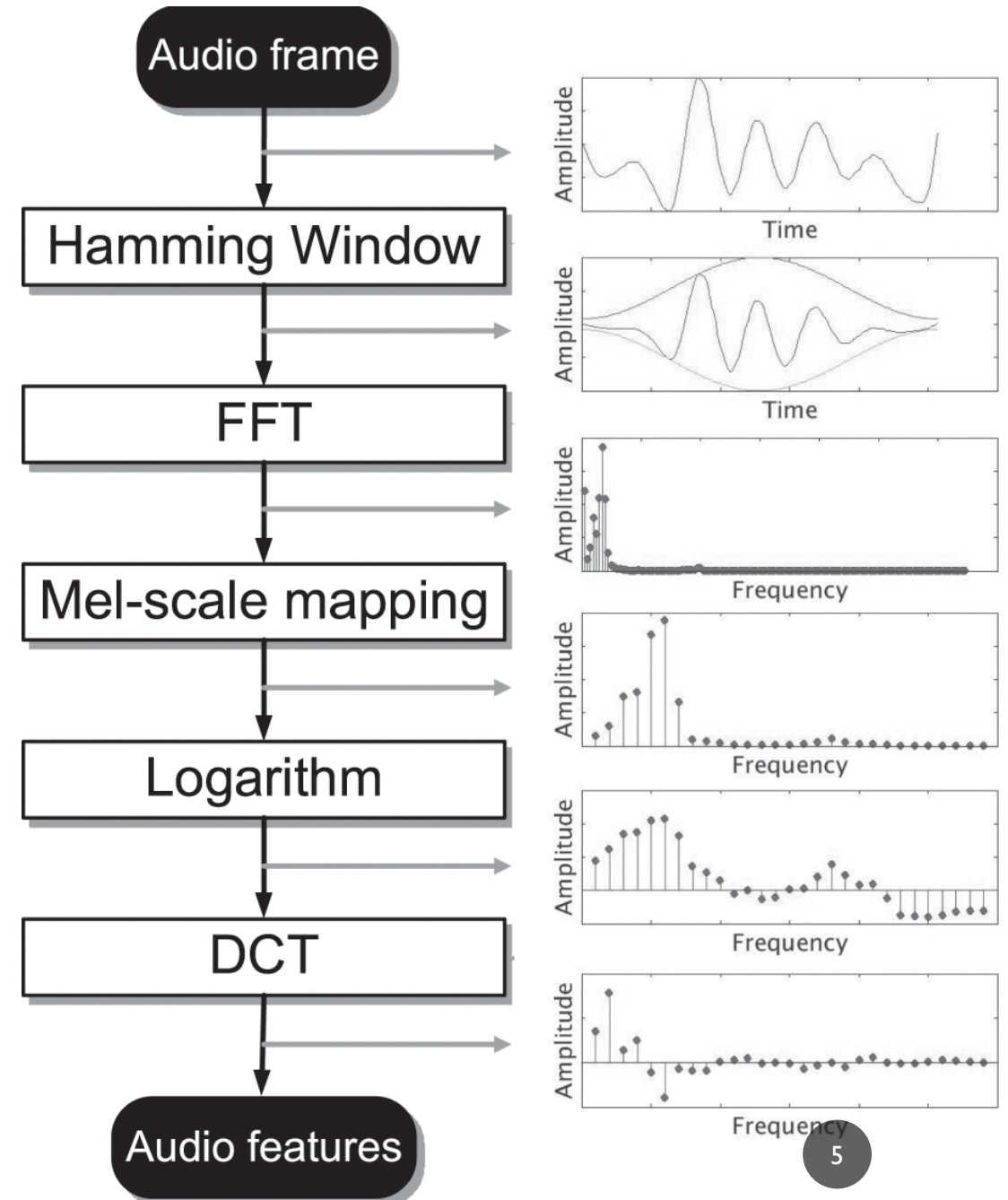
- 1000 audio – 10 genres
- 30 seconds duration



# FEATURE EXTRACTION

## MEL FREQUENCY CEPSTRAL COEFFICIENTS

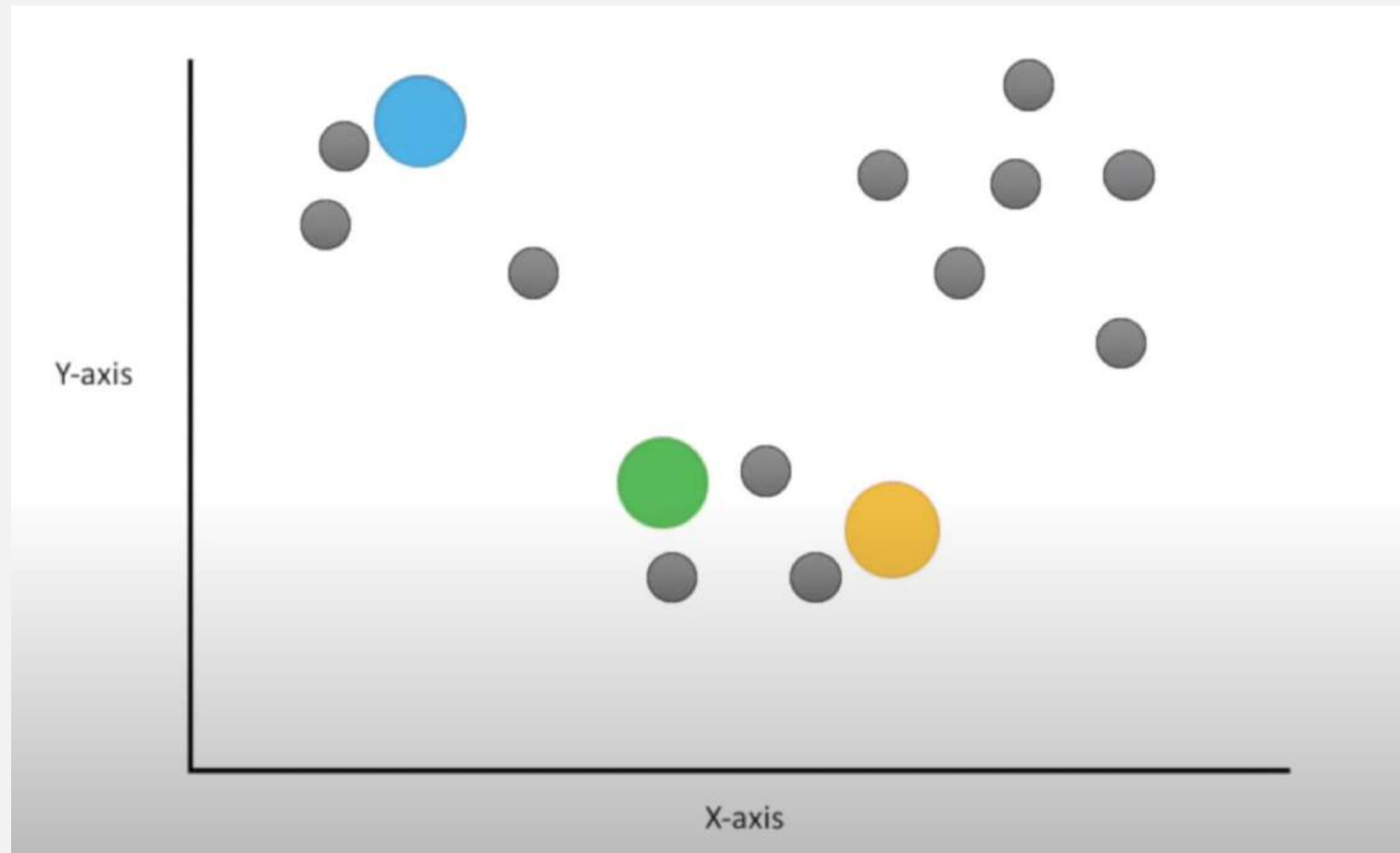
1. Divide these signals into small frames
2. Identify different frequencies present in each frame
3. Separate linguistic frequencies from the noise



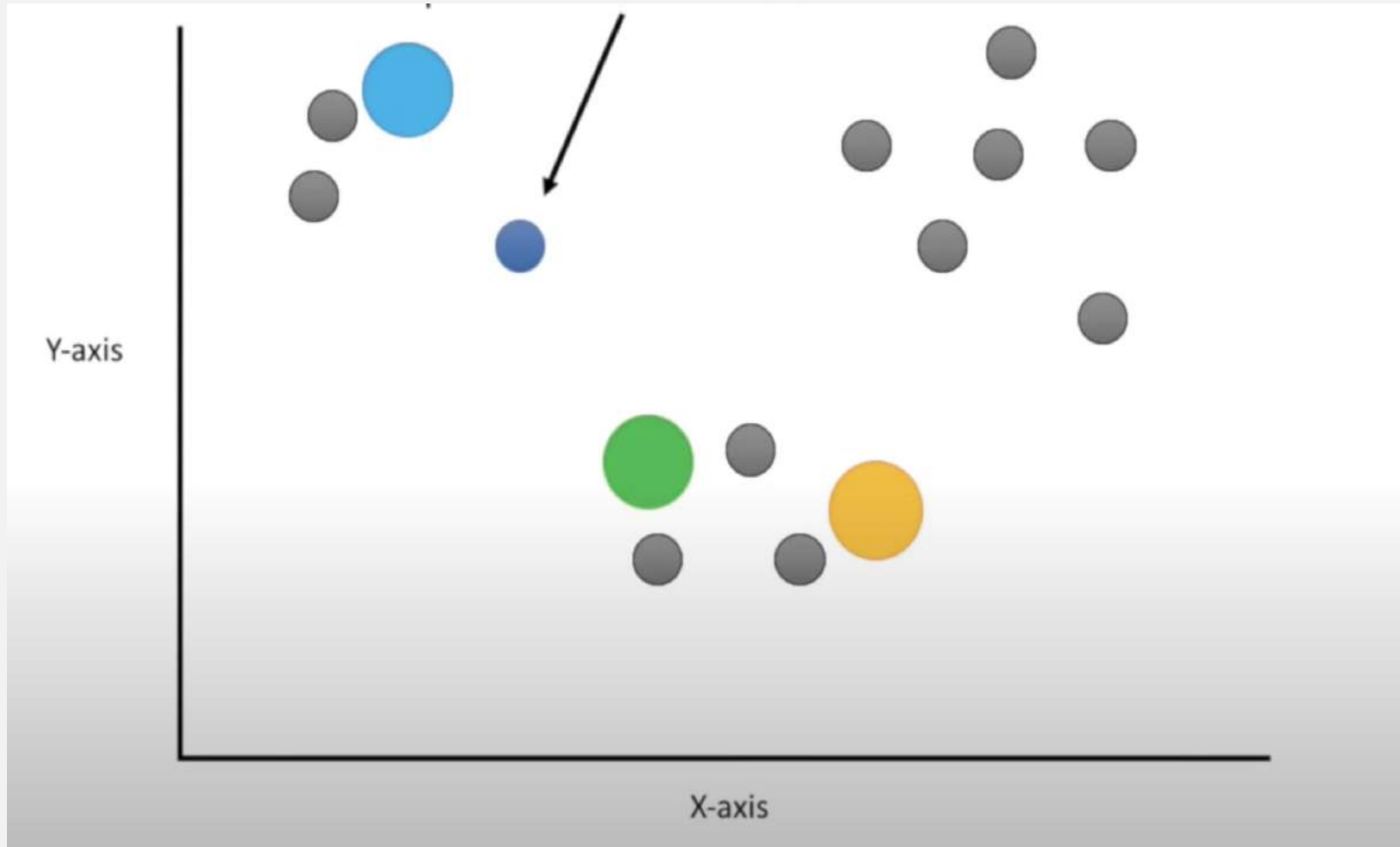
## MUSIC GENRE CLASSIFICATION APPROACH

1. K-means clustering
2. Multiclass support vector machines
3. K-nearest neighbors
4. Convolutional neural networks

# I. K-MEANS CLUSTERING

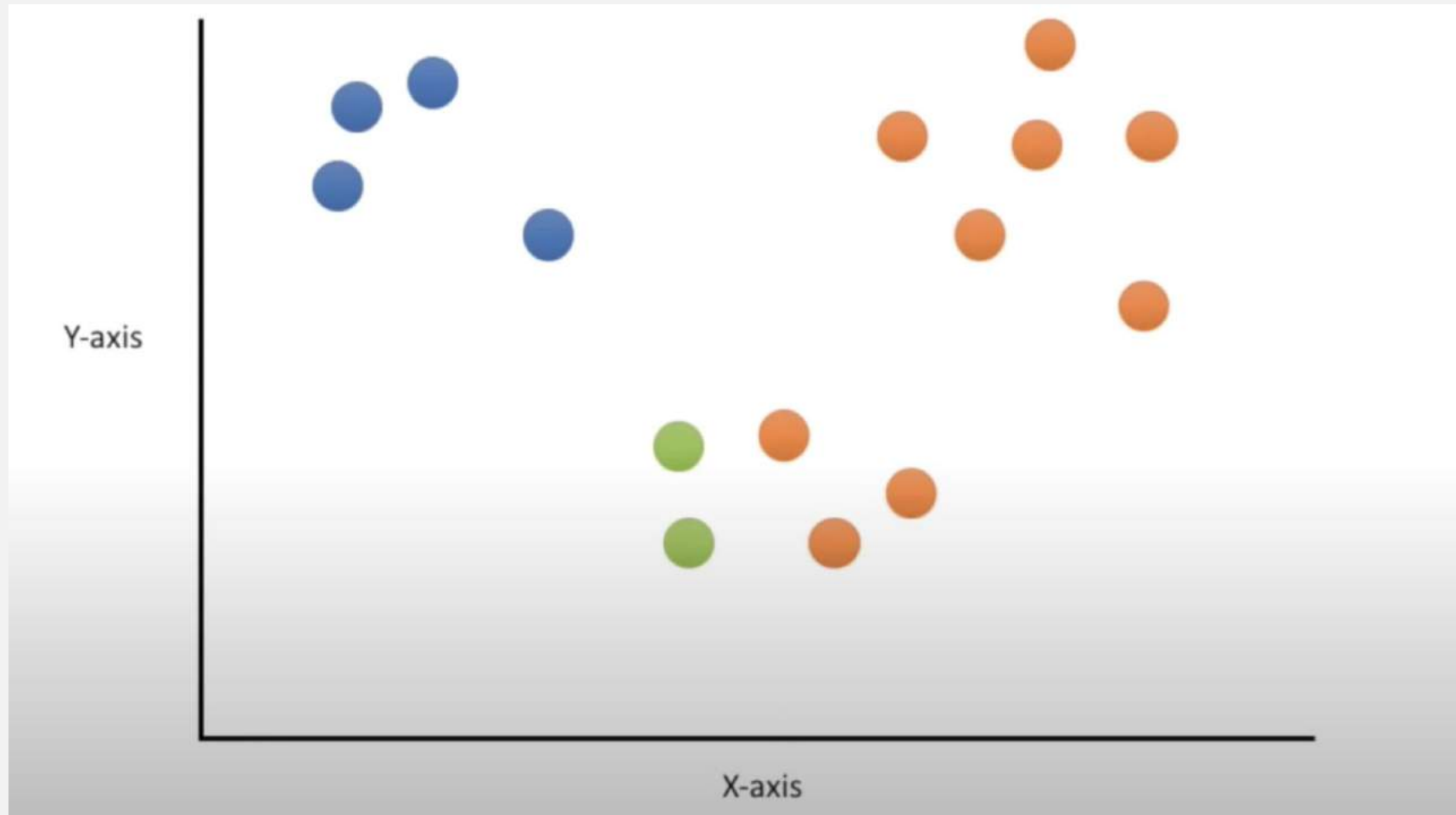


# I. K-MEANS CLUSTERING

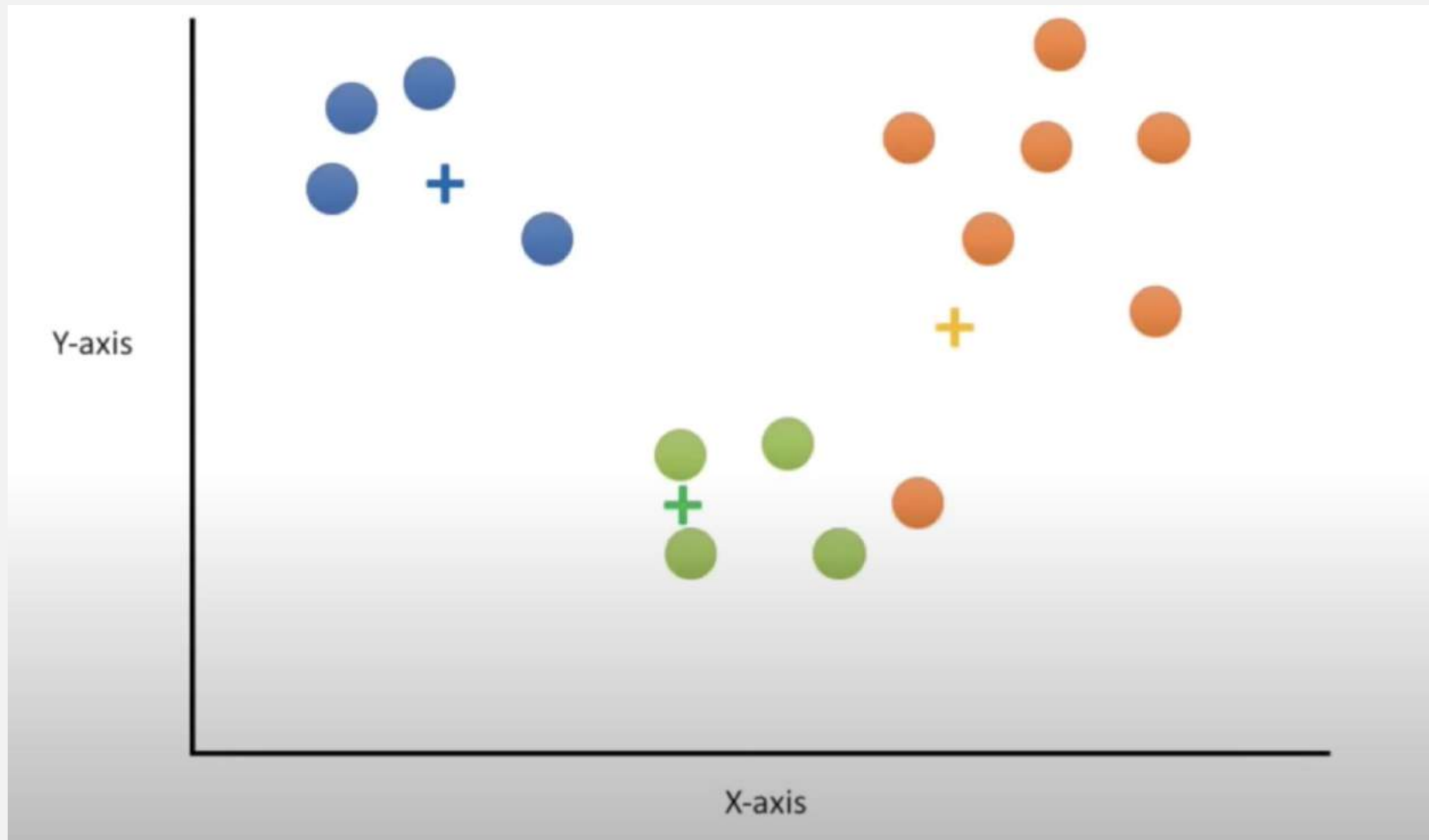




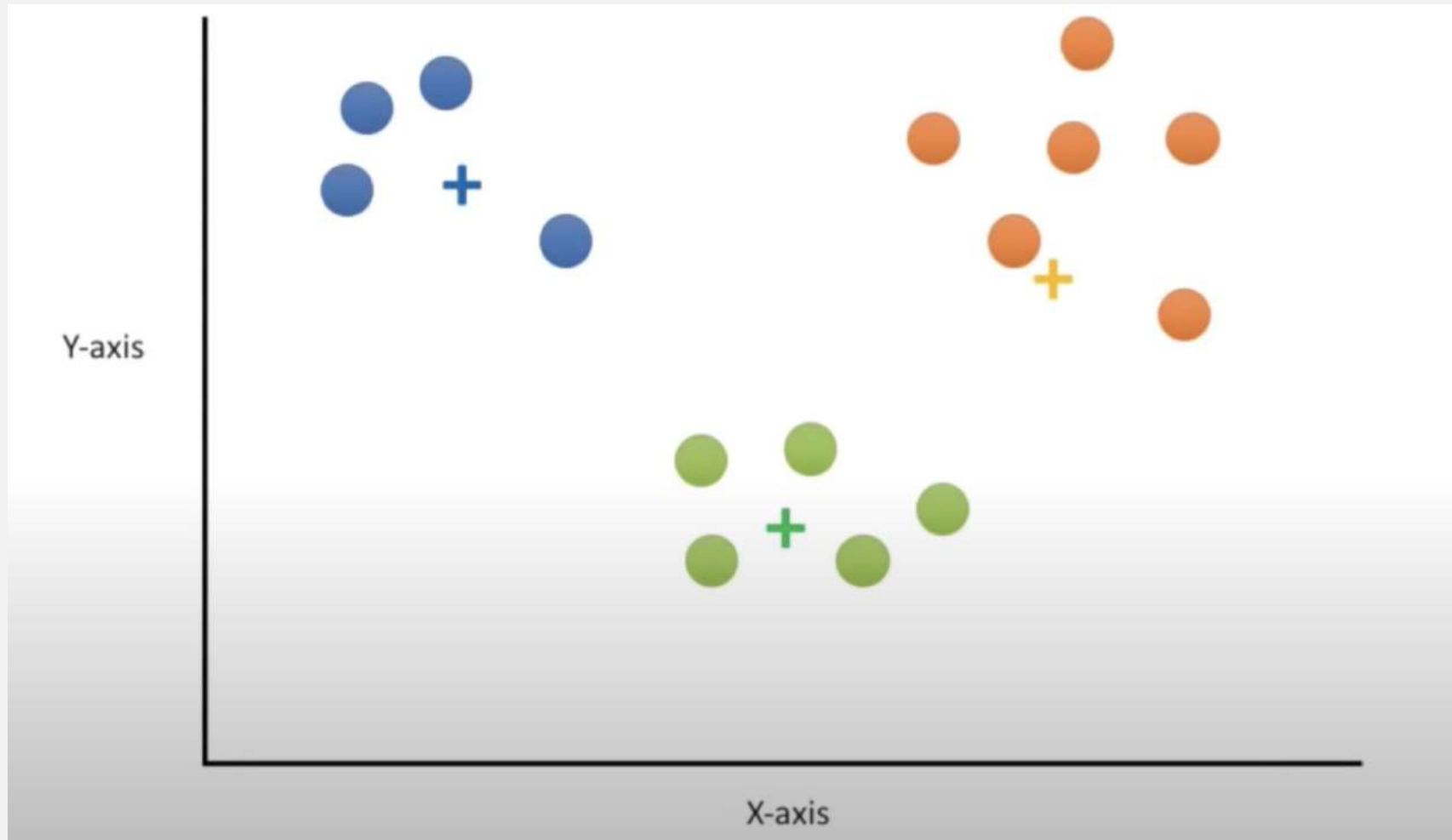
# I. K-MEANS CLUSTERING



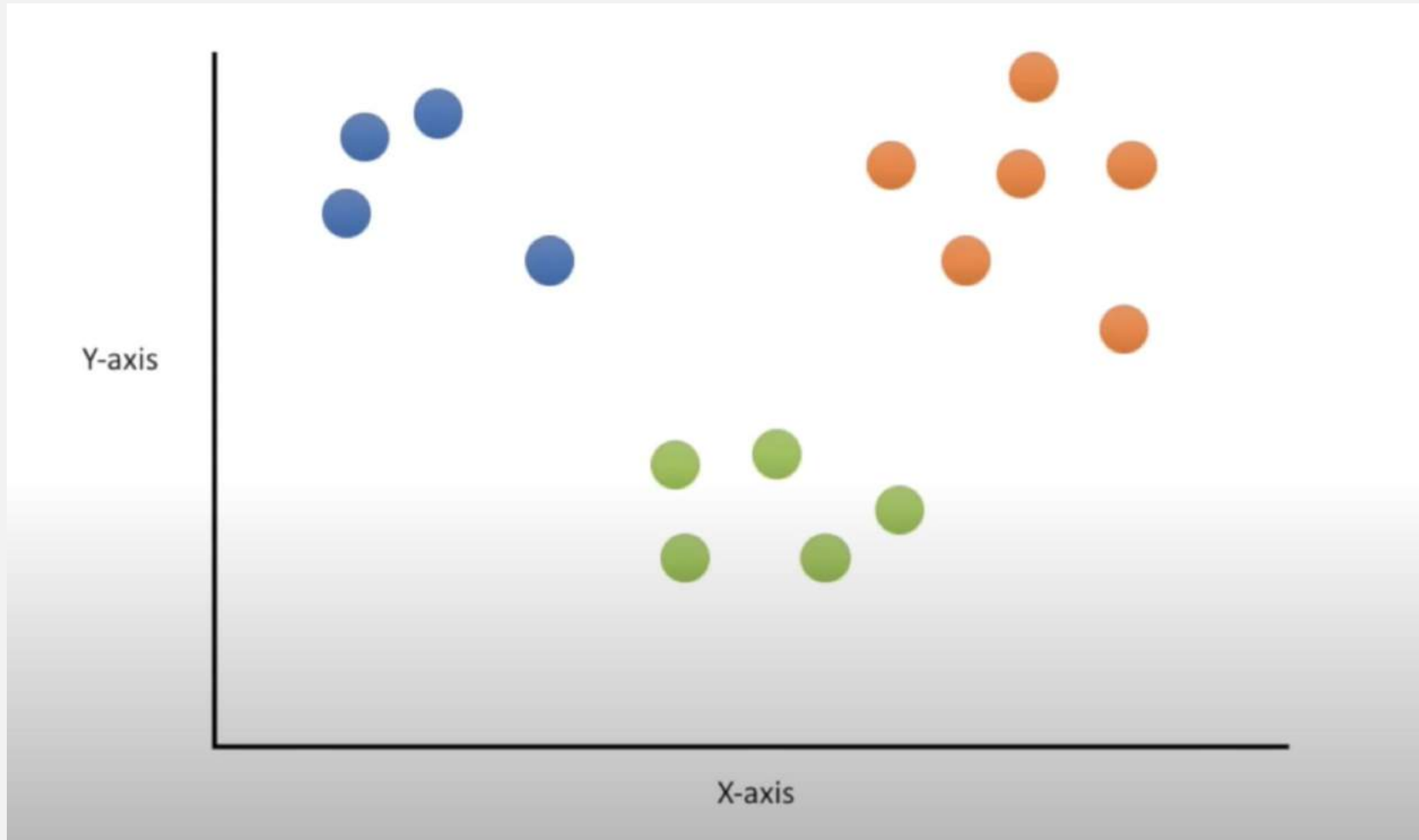
# I. K-MEANS CLUSTERING



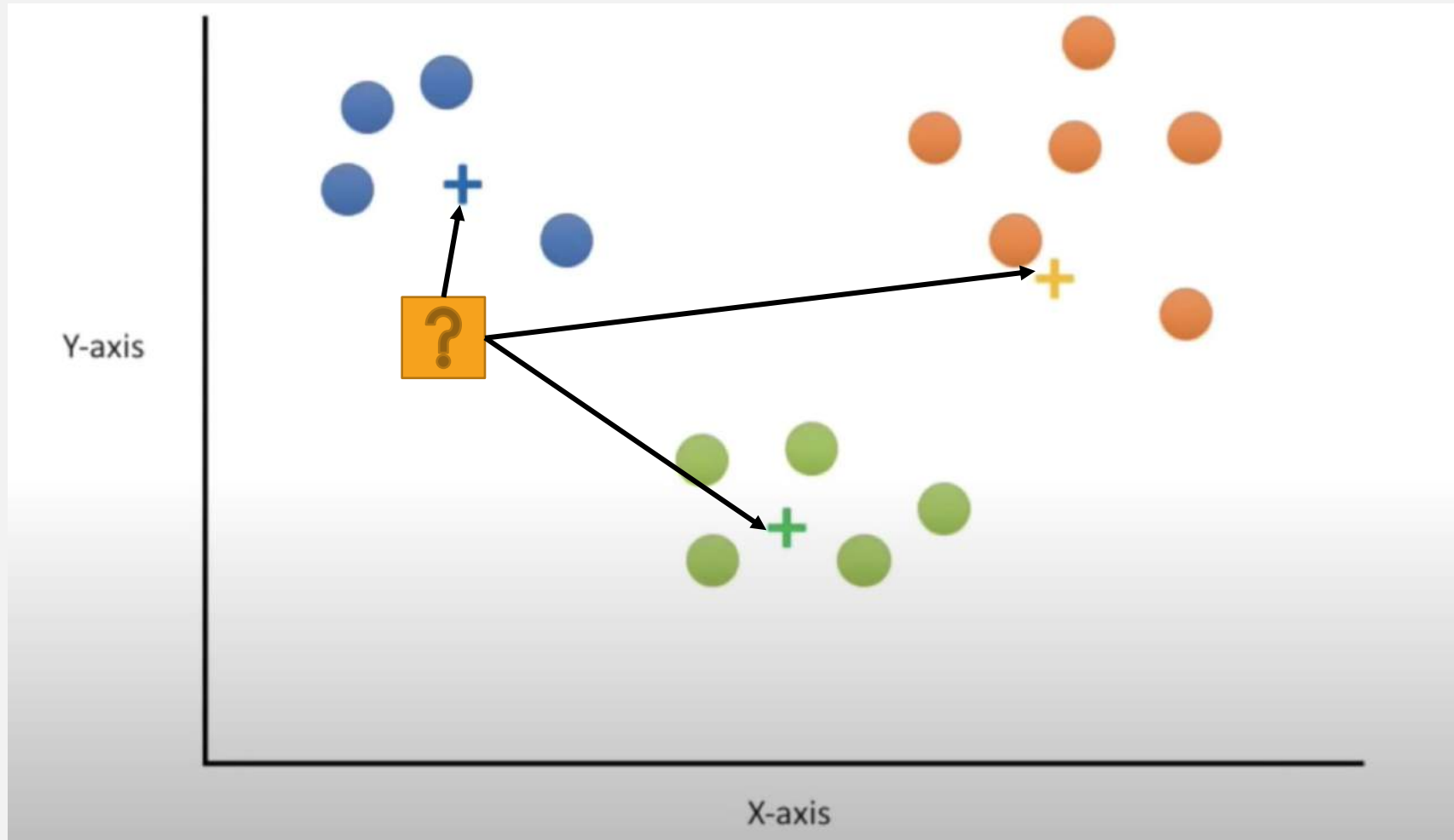
# I. K-MEANS CLUSTERING



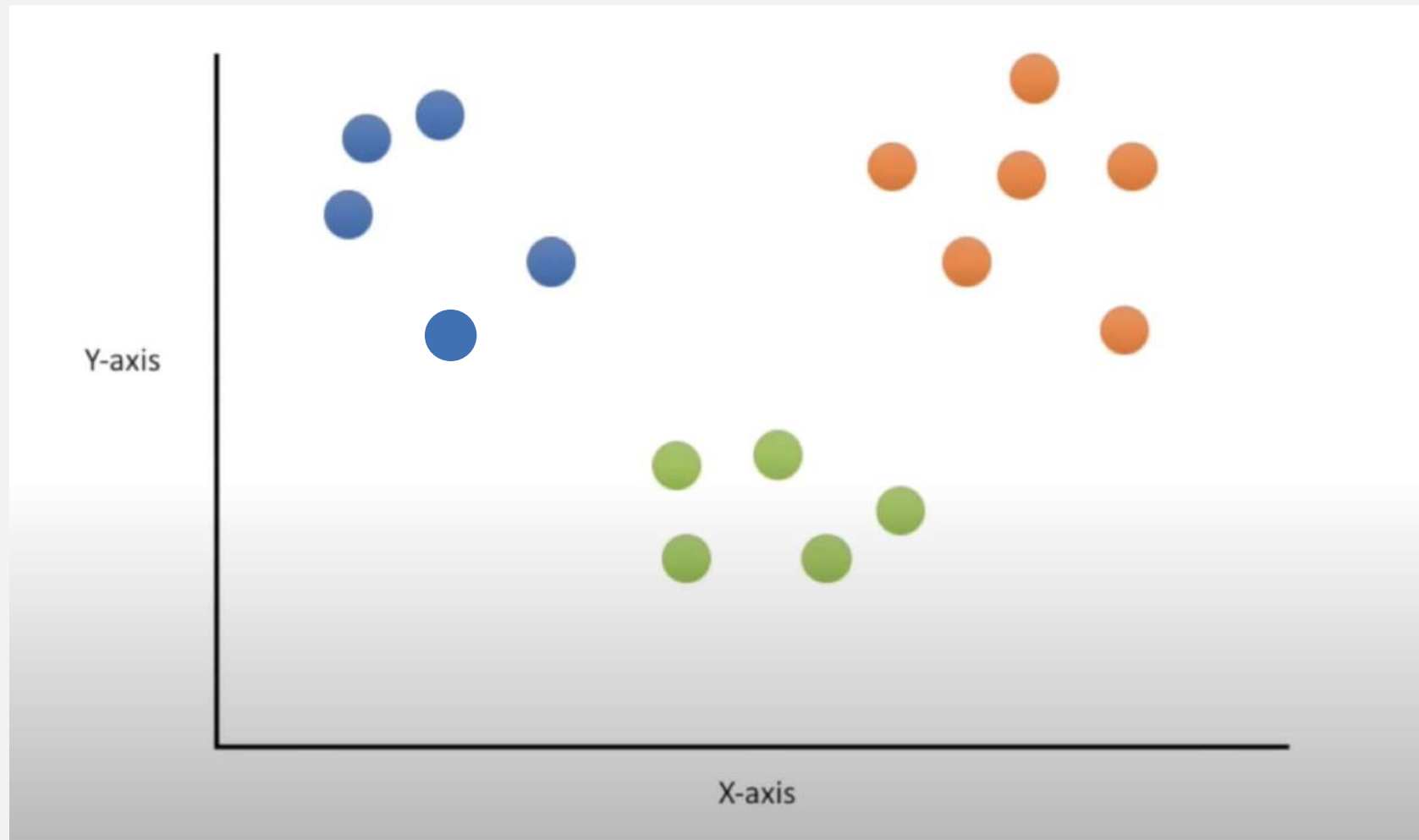
# I. K-MEANS CLUSTERING



# I. K-MEANS CLUSTERING

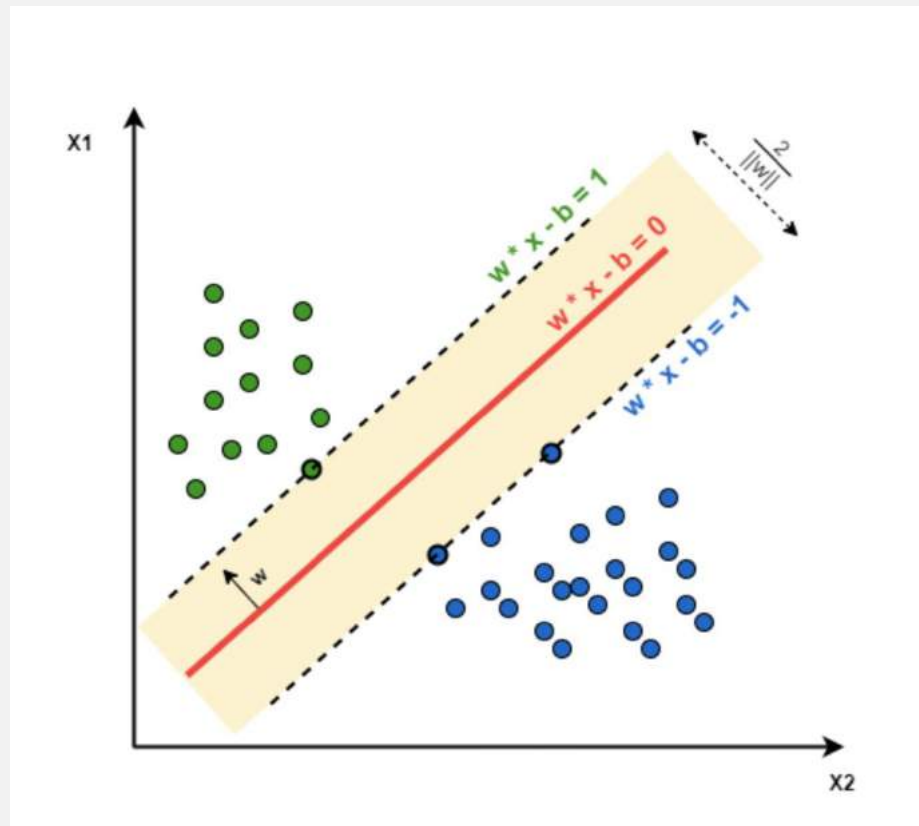


# I. K-MEANS CLUSTERING

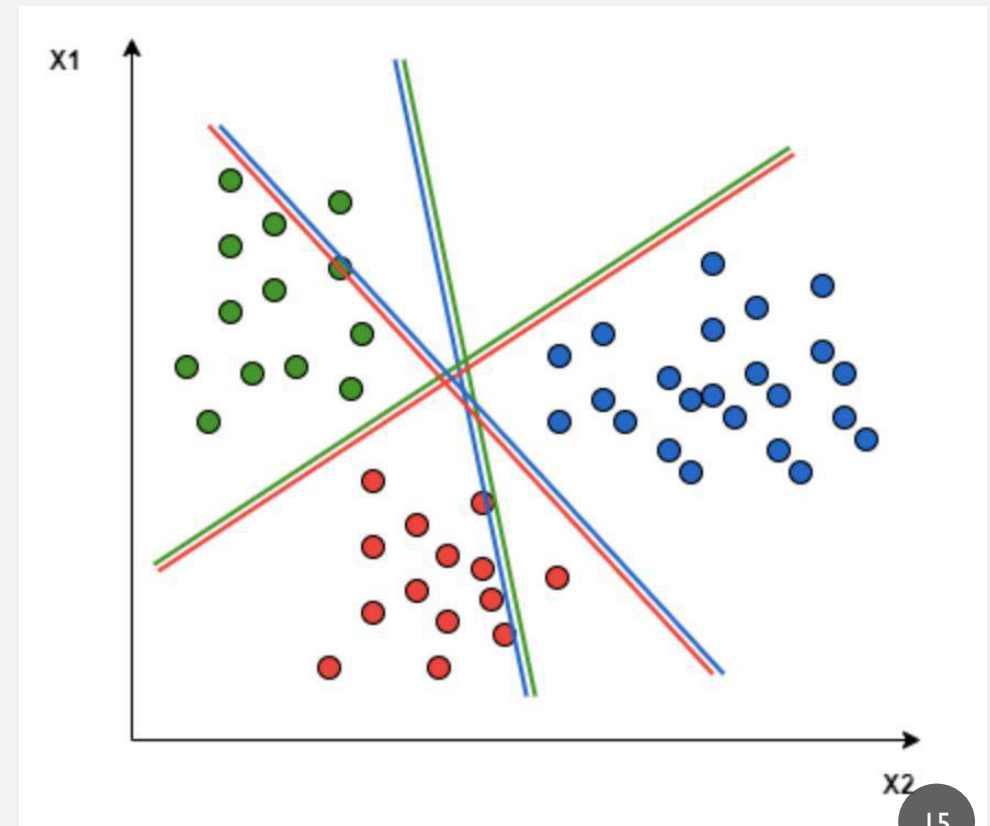


## 2. MULTICLASS SUPPORT VECTOR MACHINES

Singleclass

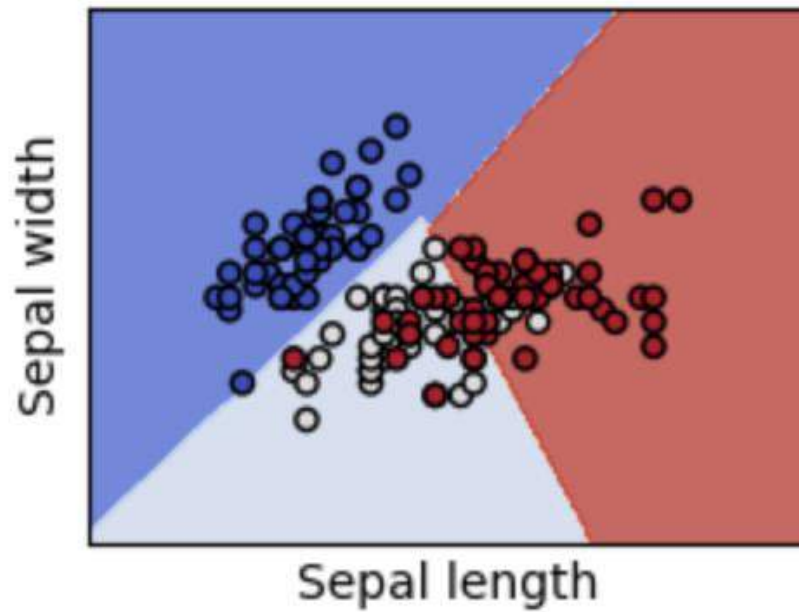


Multiclass

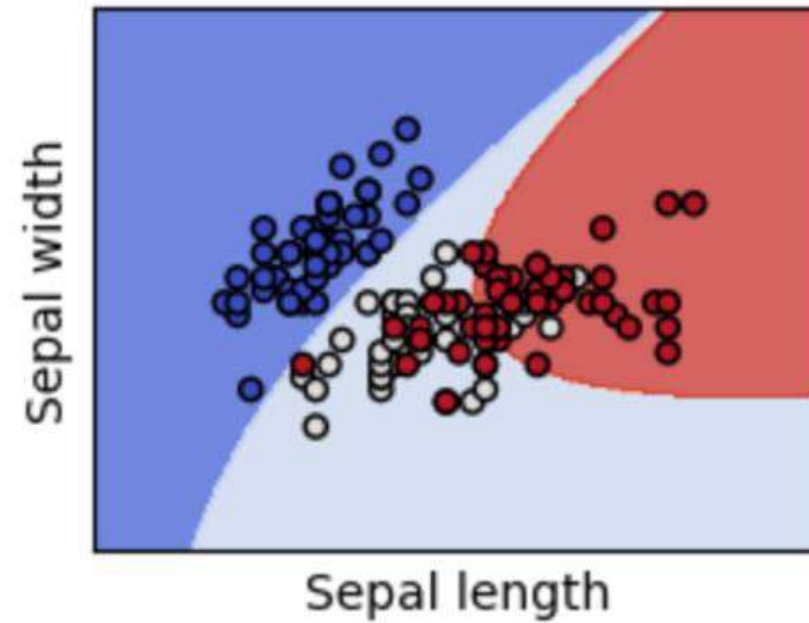


## 2. MULTICLASS SUPPORT VECTOR MACHINES

SVC with linear kernel



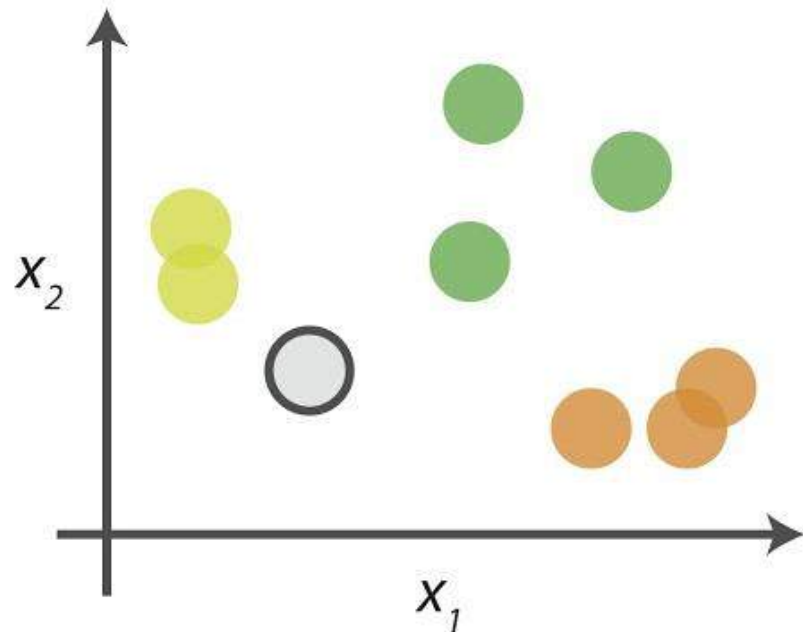
SVC with polynomial (degree 3) kernel



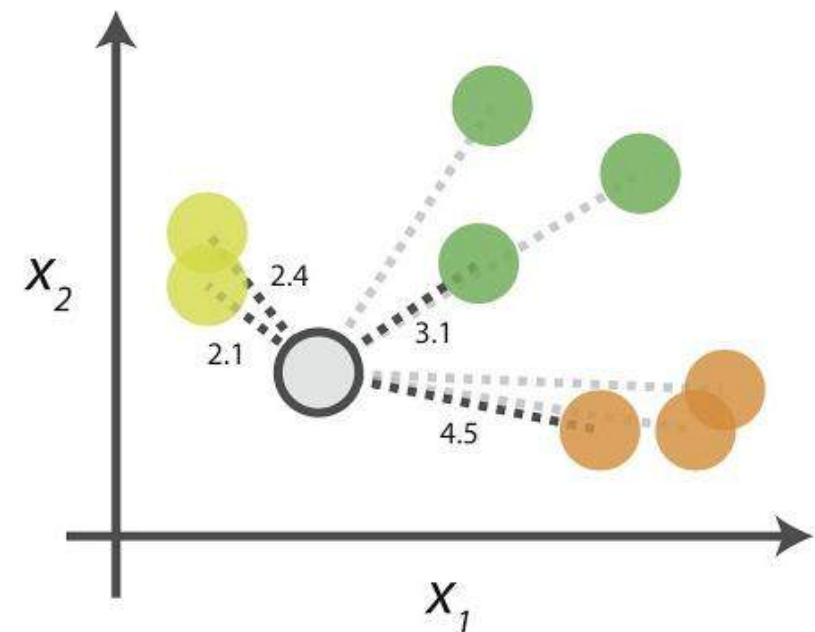


### 3. K-NEAREST NEIGHBORS

#### 0. Look at the data











#### 1. Calculate distances









## 3. K-NEAREST NEIGHBORS

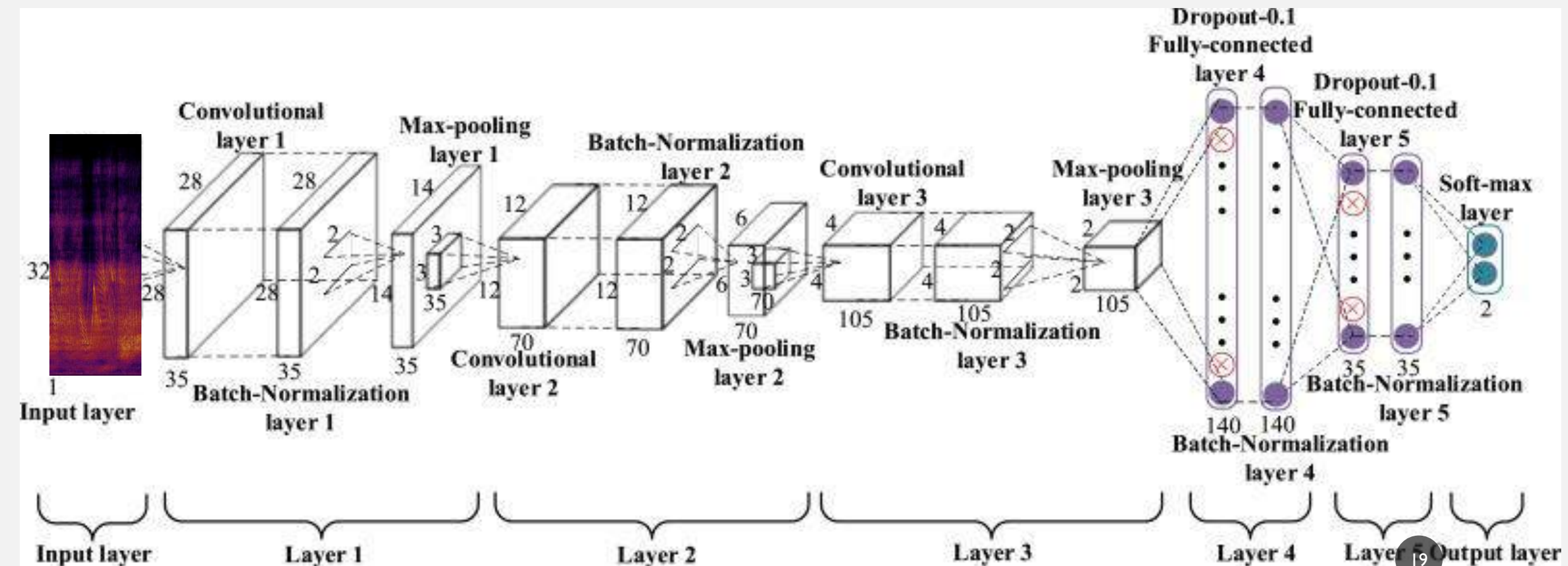
### 2. Find neighbours

<i>Point Distance</i>		
 ... 	2.1	→ 1st NN
 ... 	2.4	→ 2nd NN
 ... 	3.1	→ 3rd NN
 ... 	4.5	→ 4th NN

### 3. Vote on labels

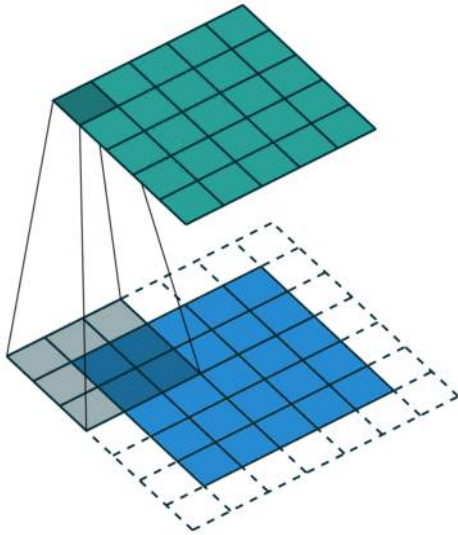
<i>Class</i>	<i># of votes</i>	
	2	→ Class  wins the vote! Point  is therefore predicted to be of class  .
	1	
	1	

## 4. CONVOLUTIONAL NEURAL NETWORKS

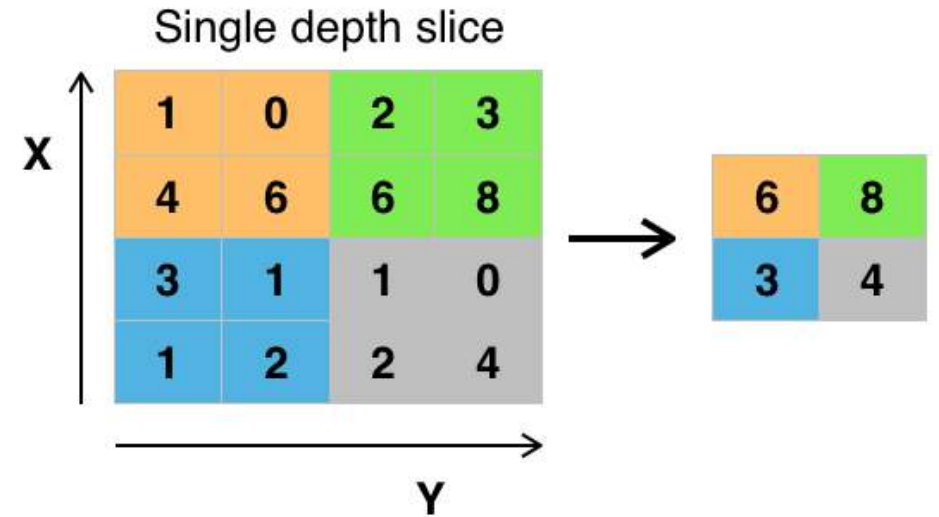


## 4. CONVOLUTIONAL NEURAL NETWORKS

Convolution Layer

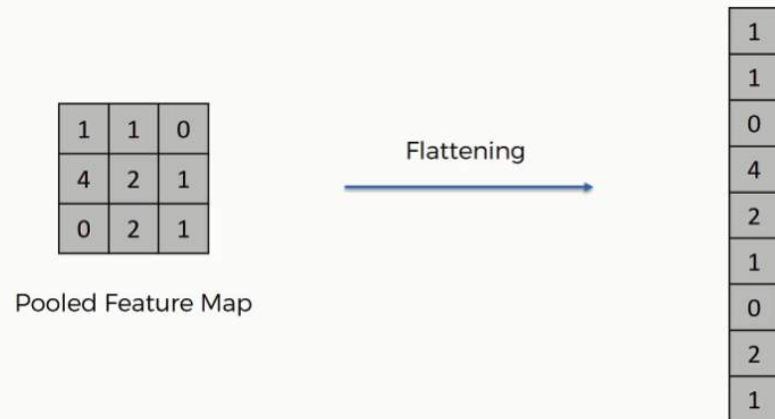


Max Pooling Layer

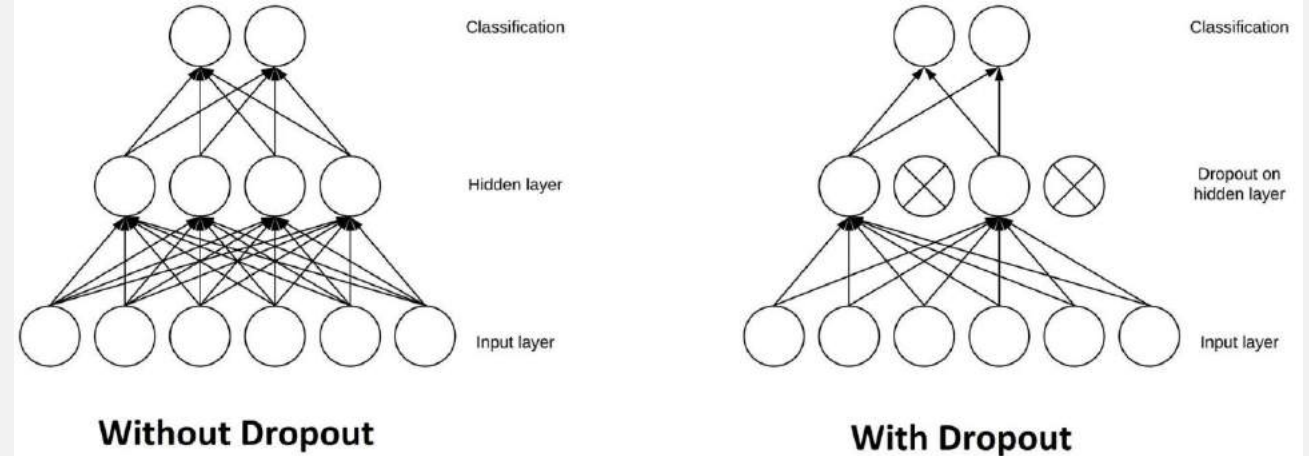


## 4. CONVOLUTIONAL NEURAL NETWORKS

### Flatten Layer



### Dropout Layer



# COMPARISON

Algorithm	Accuracy
K-Means	Inconsistent (29% - 39%)
MCSVM	Polynomial Kernel: 66.00% Linear Kernel: 69.50%
KNN	67.34%
CNN	72.03%

THANK YOU FOR LISTENING