



AI/MACHINE LEARNING WORKSHOP

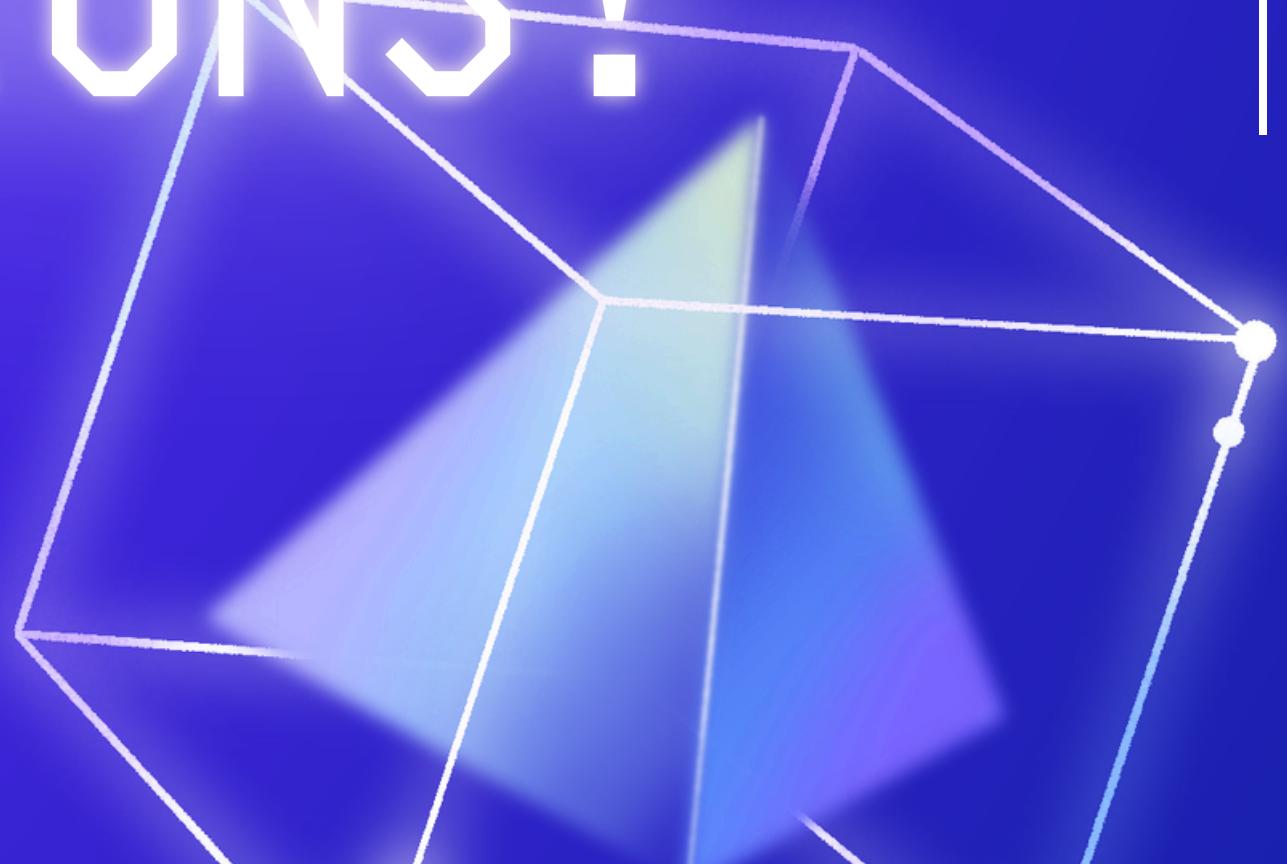
DAY 8: INTRODUCTION TO AI & ML: K-NEAREST NEIGHBOR

Youth Opportunities in Tech Innovation





REMINDER PLEASE
ASK QUESTIONS!



WHAT IS K- NEAREST NEIGHBOR

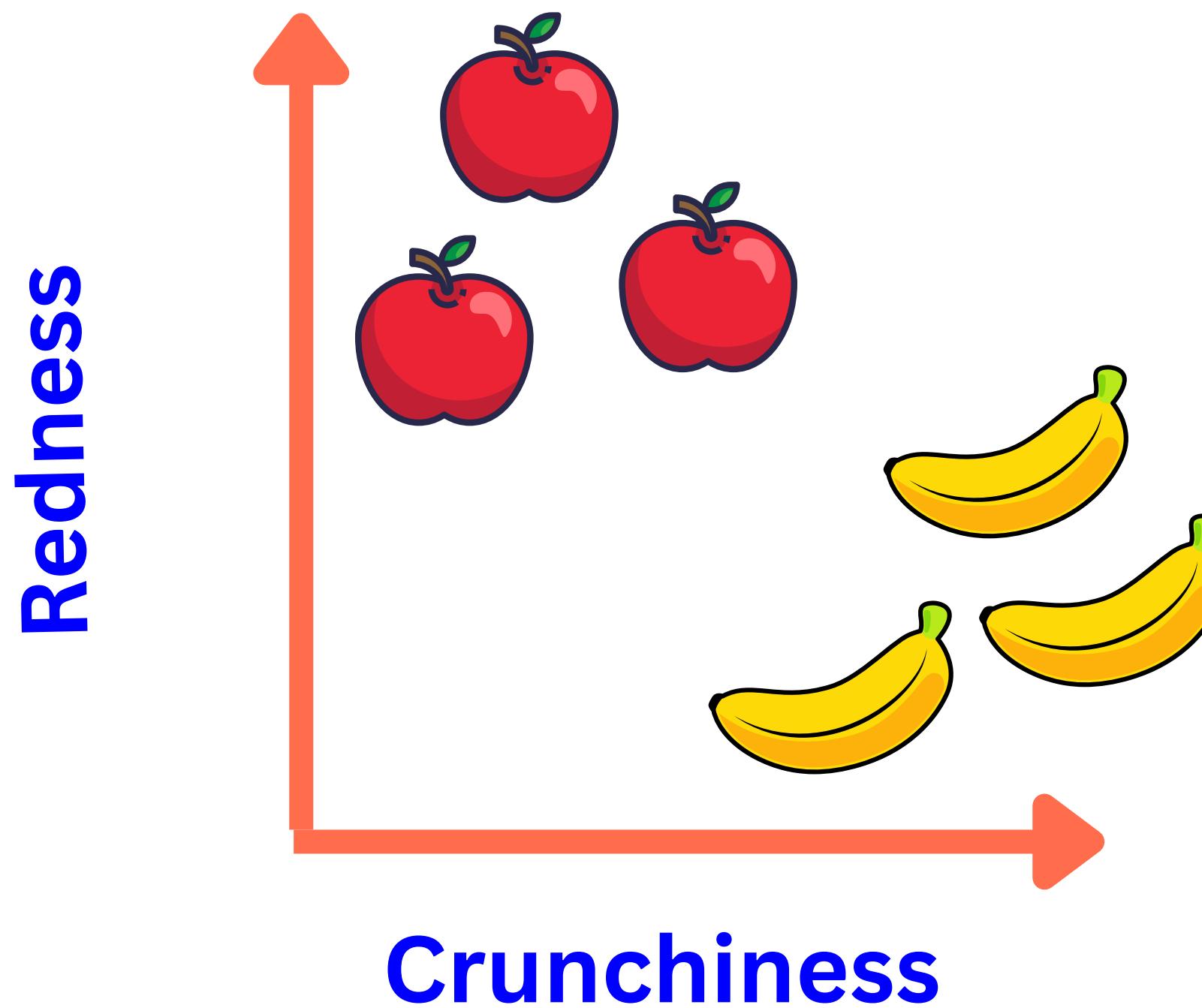
SUPERVISED MACHINE LEARNING ALGORITHM
THAT CLASSIFIES OR PREDICTS THE VALUE
OF A NEW DATA POINT BASED ON THE
MAJORITY CLASS (OR AVERAGE VALUE) OF
ITS "K" NEAREST NEIGHBORS IN THE
TRAINING DATA



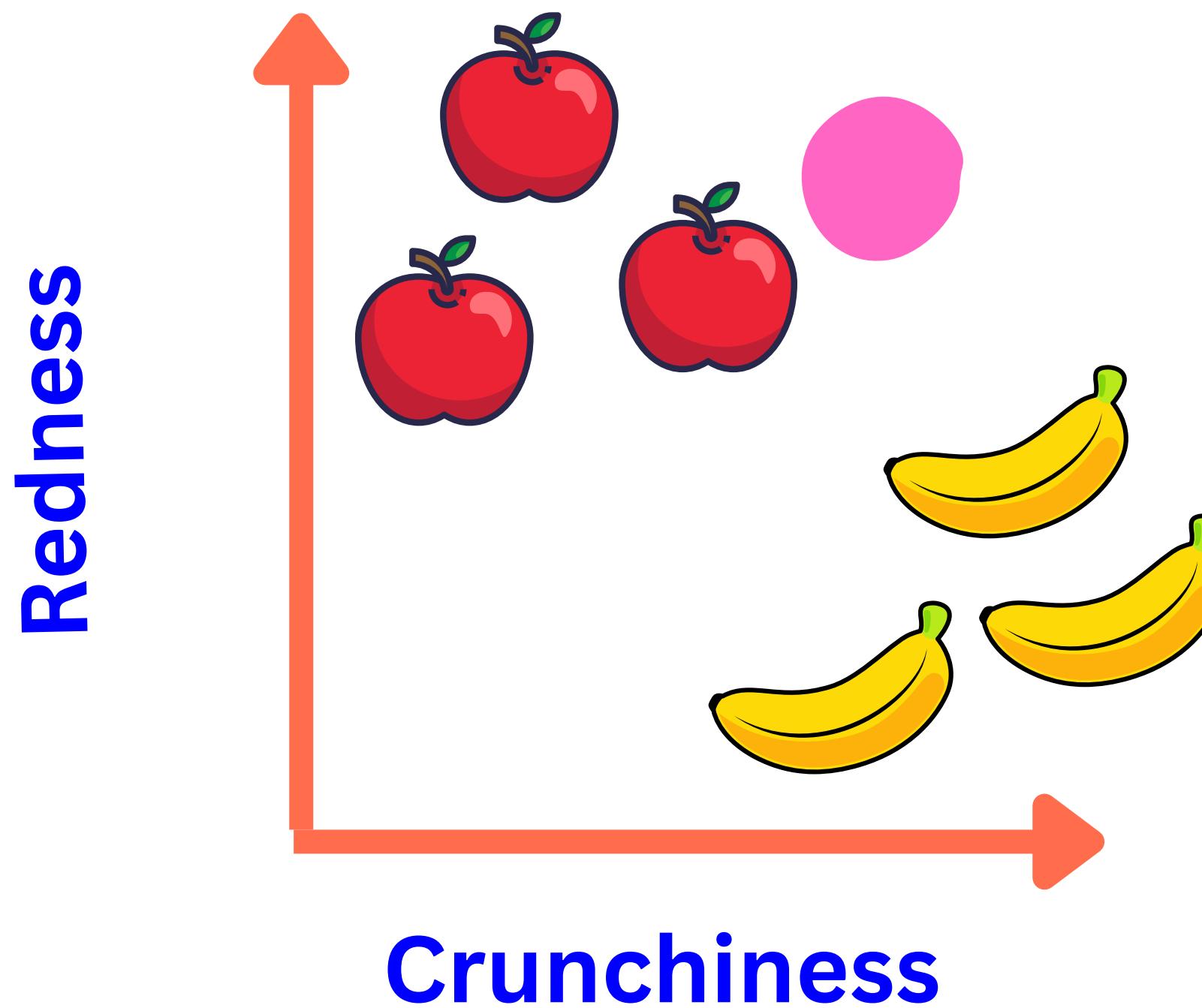
Real world example



KNN Model

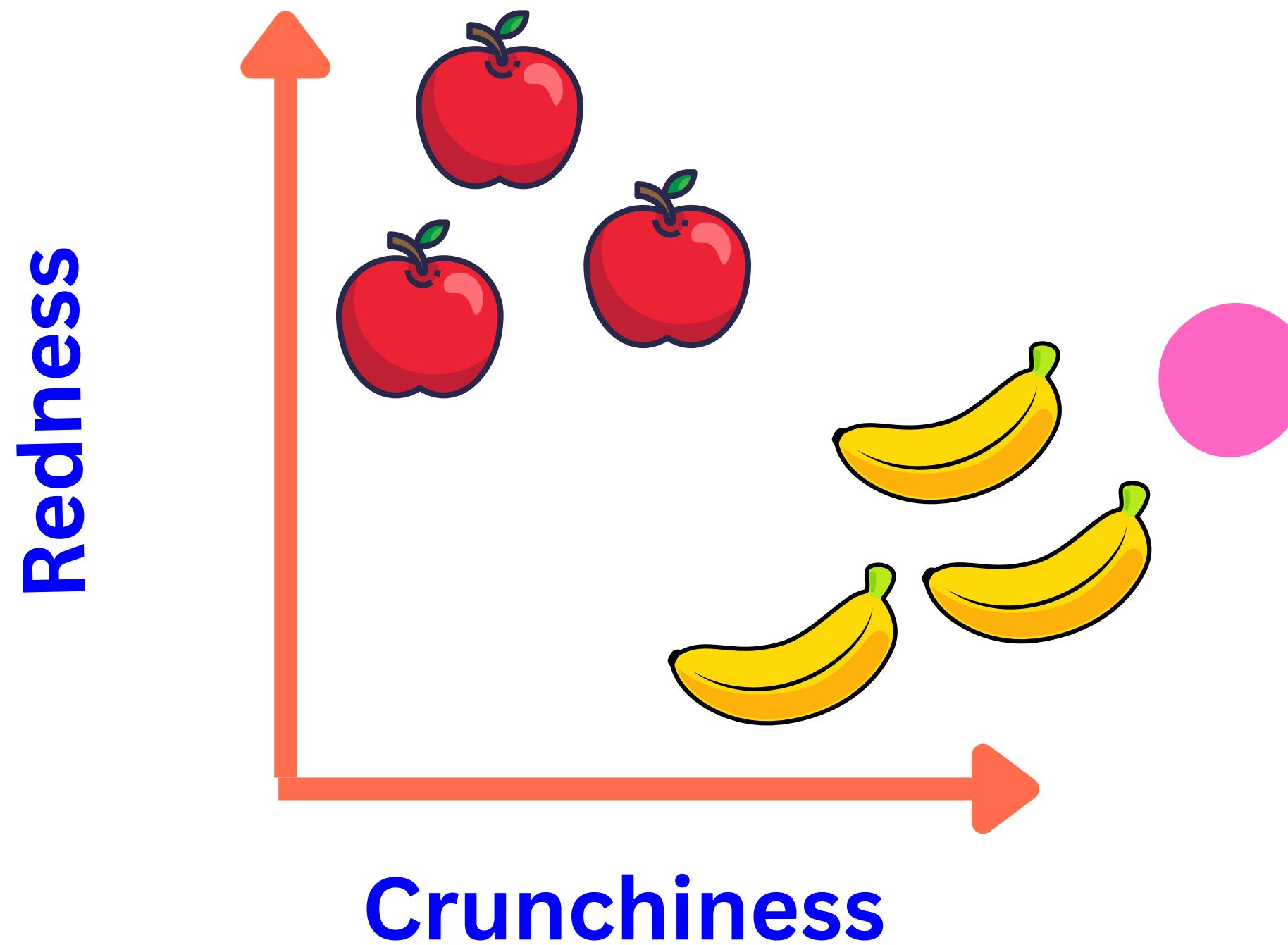


KNN Model



KNN Model

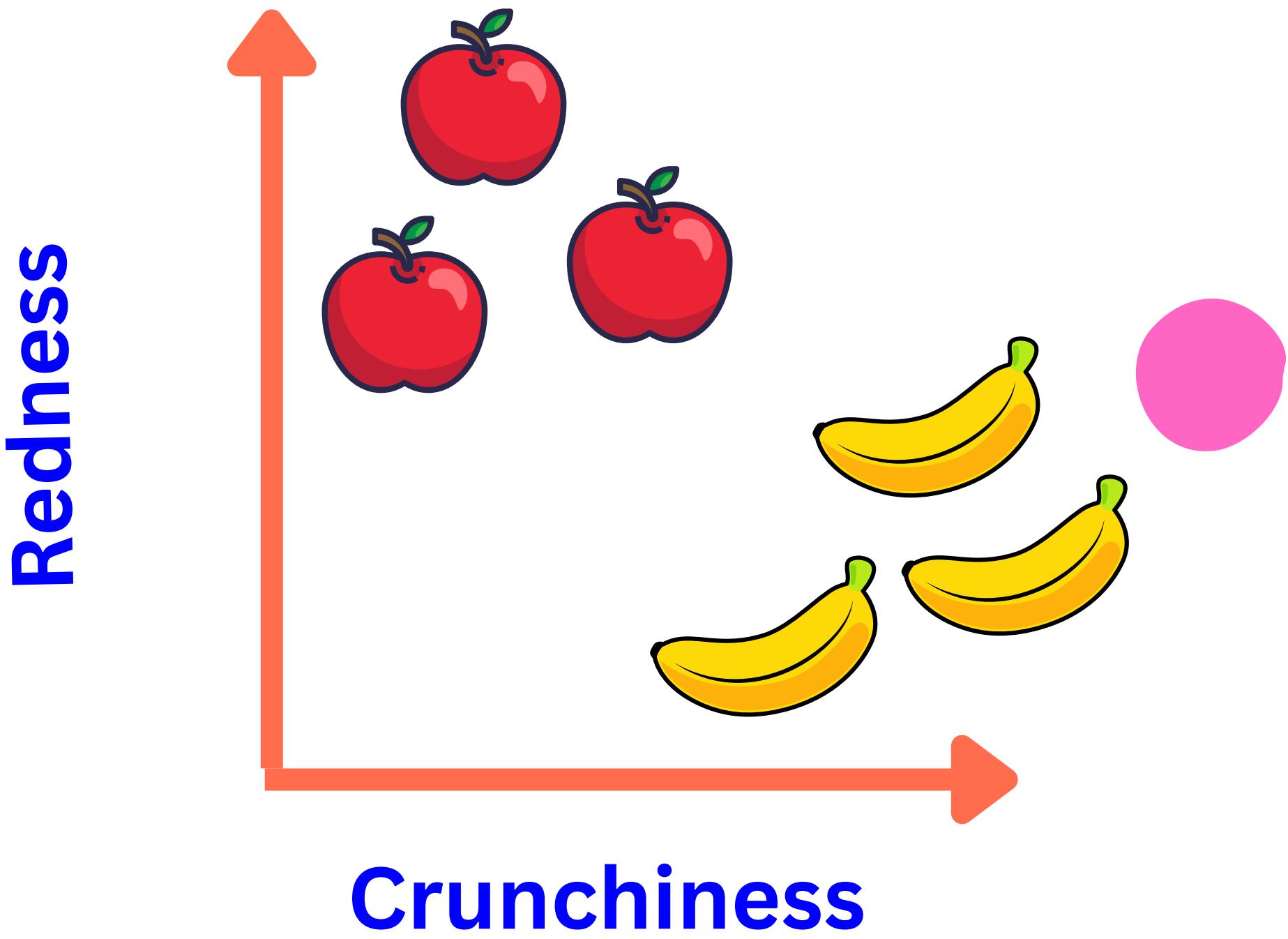
Similar data points are located near each other



KNN Model

Step-by-Step: How KNN Works

- Choose a value for k (number of neighbors).
- Calculate the distance from the new point to all other points in the dataset.
- Sort distances and pick the k nearest points.
- Vote (classification) or average (regression) to make a prediction.



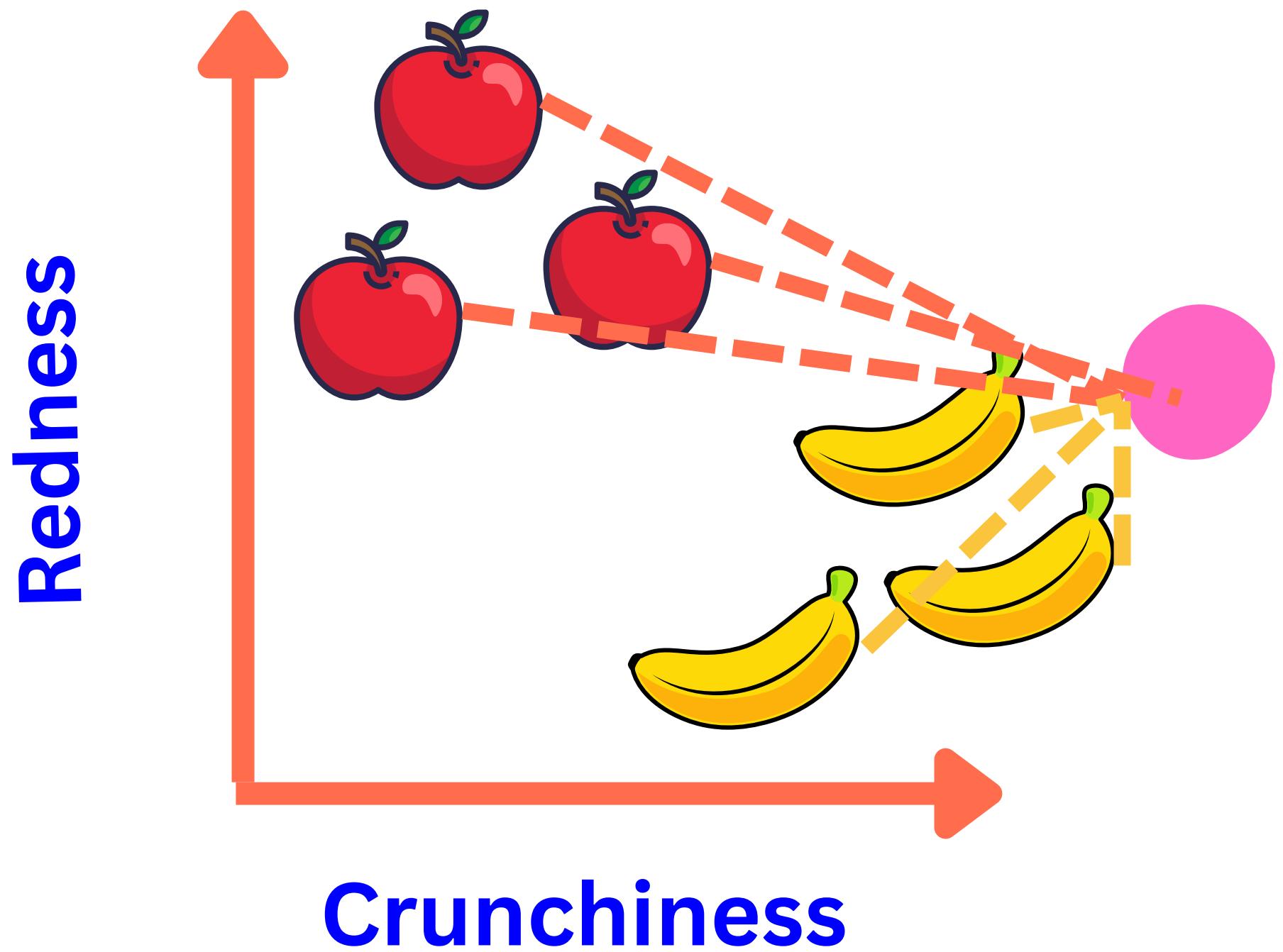
KNN Model

Calculate the Distance:

Euclidean Distance

- New input: $\vec{x} = (x_1, x_2, \dots, x_n)$
- Training point: $\vec{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)})$

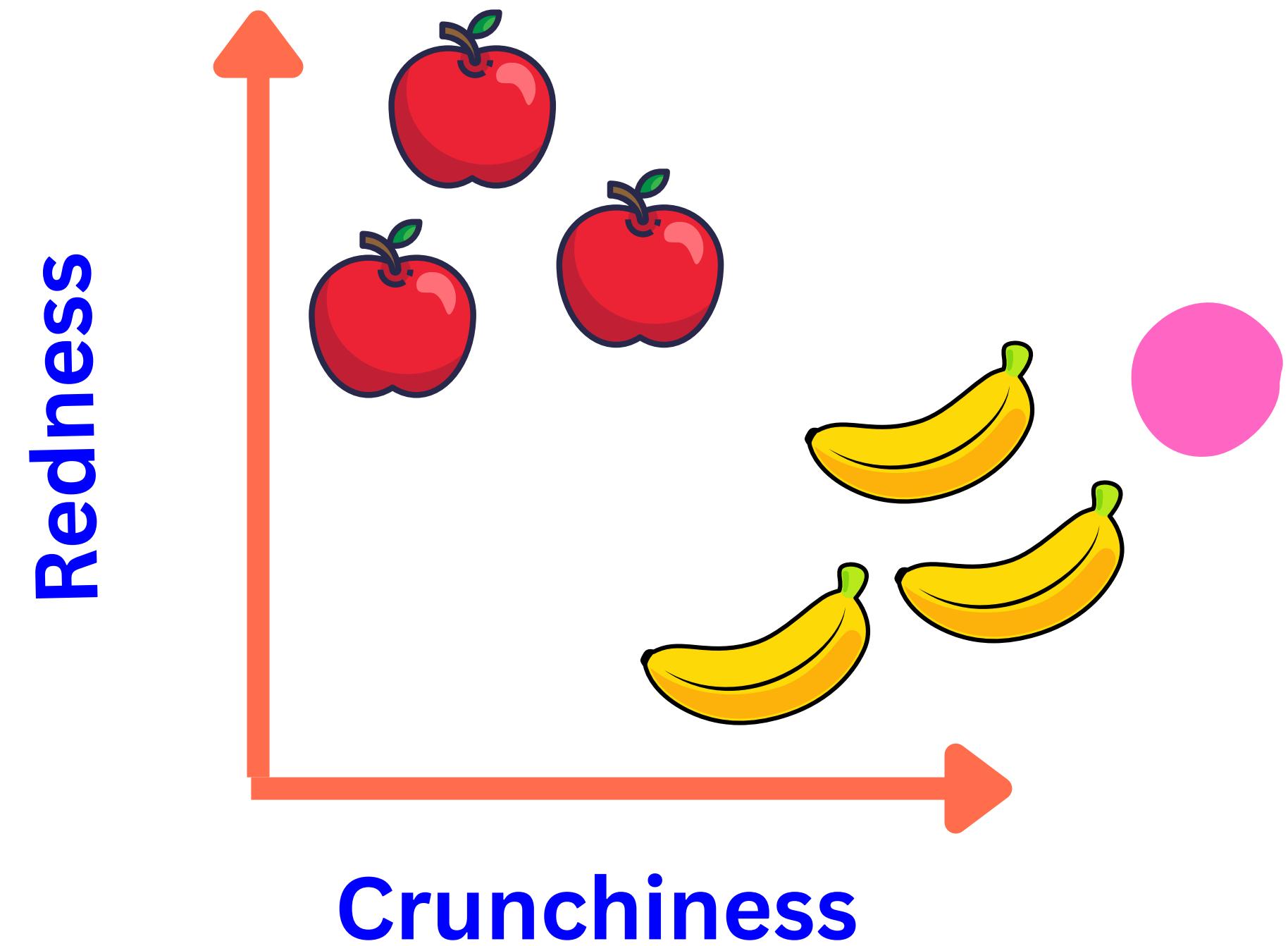
$$d(\vec{x}, \vec{x}^{(i)}) = \sqrt{\sum_{j=1}^n (x_j - x_j^{(i)})^2}$$



KNN Model

Sort the Distances

D1	[Yellow]	[Yellow]	[Yellow]
D2	[Yellow]	[Yellow]	[Yellow]
D3	[Yellow]	[Yellow]	[Yellow]
D4	[Orange]	[Orange]	[Orange]
D5	[Orange]	[Orange]	[Orange]
D6	[Orange]	[Orange]	[Orange]



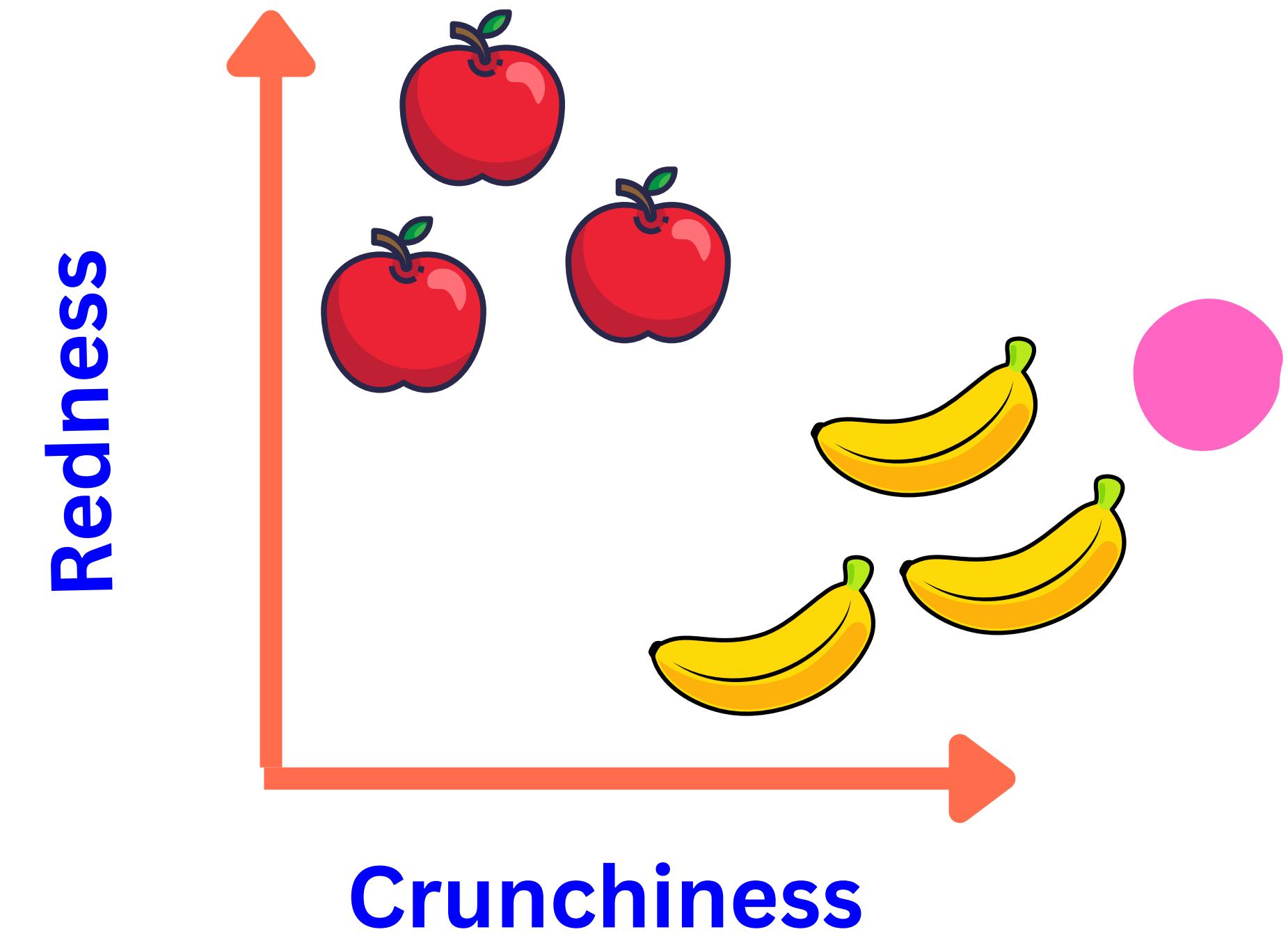
KNN Model

Sort the Distances

D1	[Yellow]	[Yellow]	[Yellow]
D2	[Yellow]	[Yellow]	[Yellow]
D3	[Yellow]	[Yellow]	[Yellow]
D4	[Red]	[Red]	[Red]
D5	[Red]	[Red]	[Red]
D6	[Red]	[Red]	[Red]

Majority Voting System

$$\hat{y} = \text{mode}(y^{(i)} \text{ for } i \in N_k)$$



KNN Model

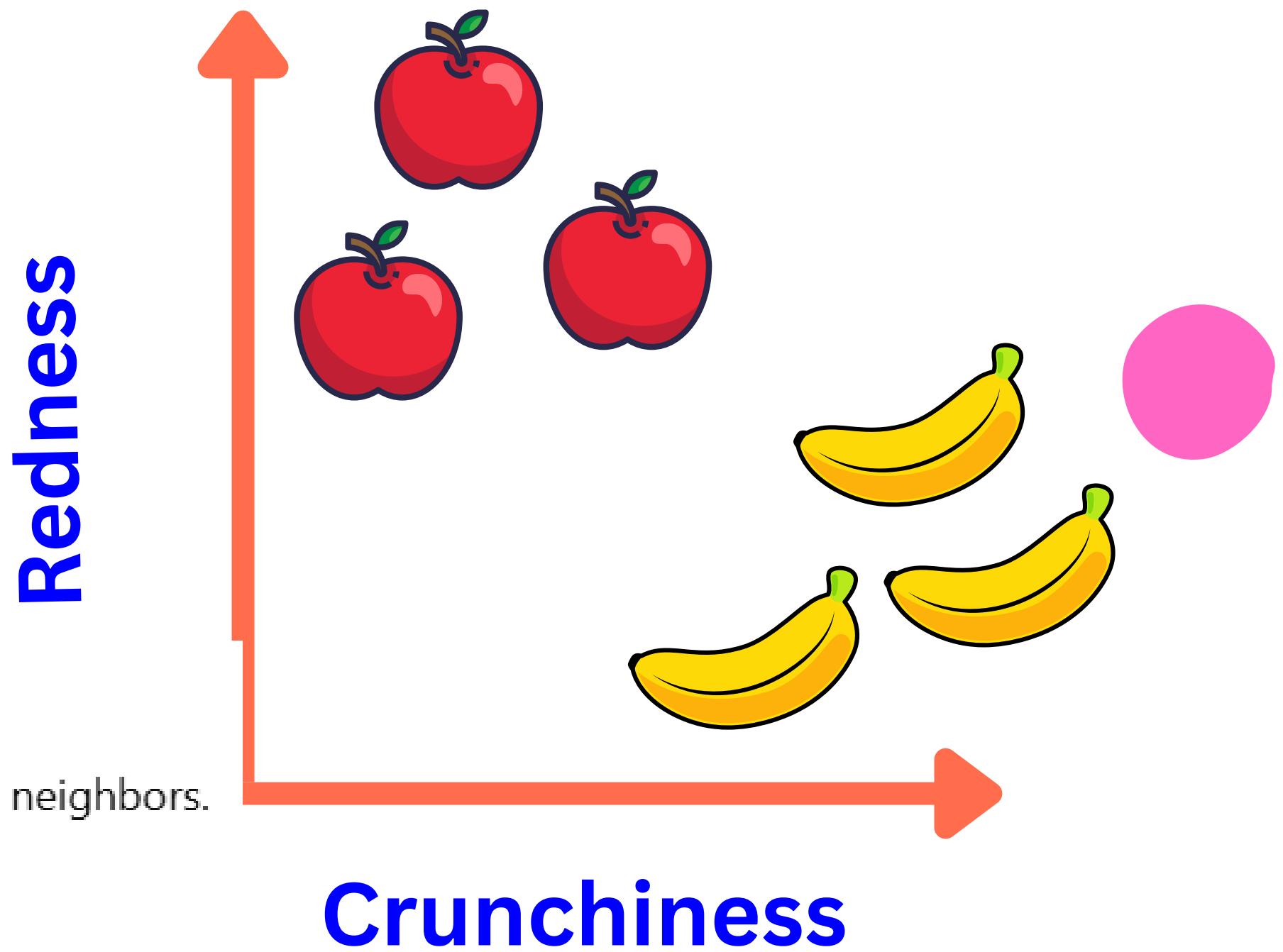
Example

- Class 1: 3 times
- Class 0: 2 times

$$\hat{y} = \text{mode}([1, 0, 1, 1, 0]) = 1$$

Prediction = 1

You're predicting the class with the highest frequency among the neighbors.



QUESTIONS AND
FINAL
THOUGHTS!

THANK YOU!