What is a Support Vector Machine (SVM)?

What is a Support Vector Machine (SVM)?

- Imagine you have two types of fruits: *apples and oranges* mixed together on a table.

- You want to draw ***a straight line*** on the table so all apples are on one side and all oranges on the other.

- That is what a SVM does: *it draws the best straight line to separate two groups*.

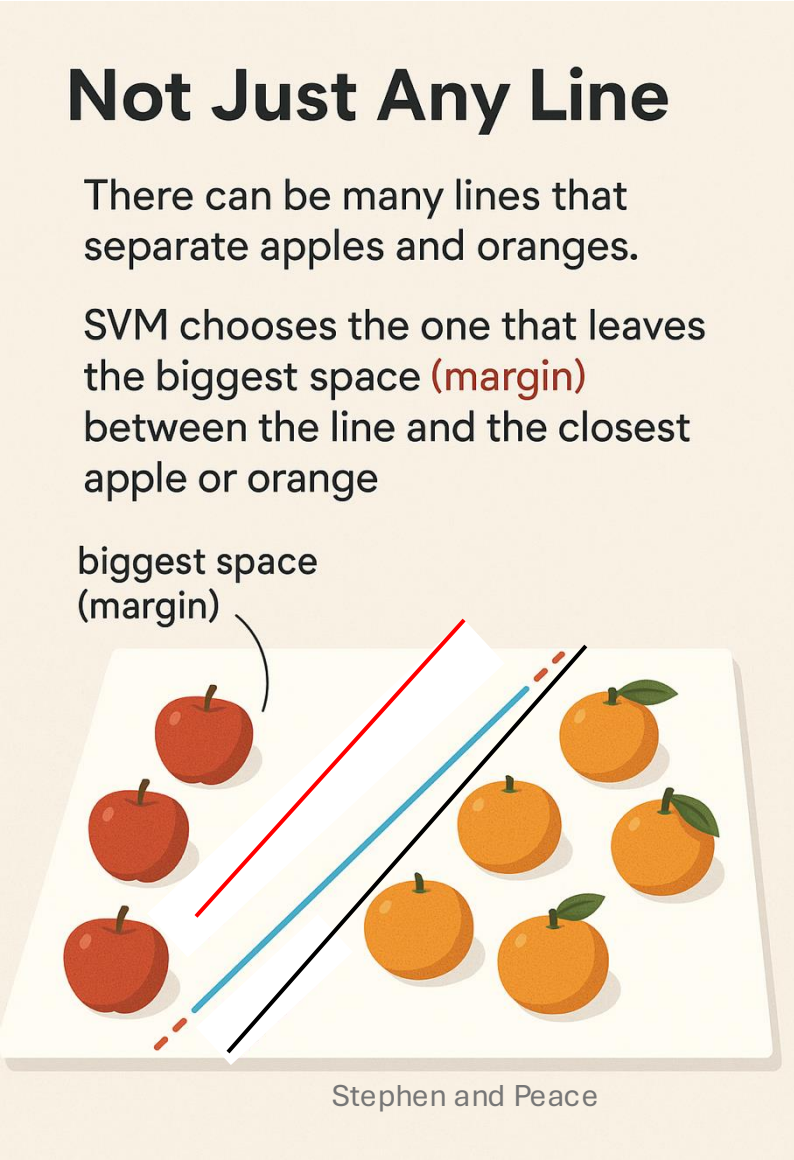
Why “Best” Line?

- *Not Just Any Line*
- There can be many lines that separate apples and oranges.

- SVM chooses the one that **leaves the biggest space** (margin) between the line and the closest apple or orange.

- **Bigger space** = more confidence the line will still work for new fruits.

Why “Best” Line?



- *Blue line is the best that can* separate apples and oranges.

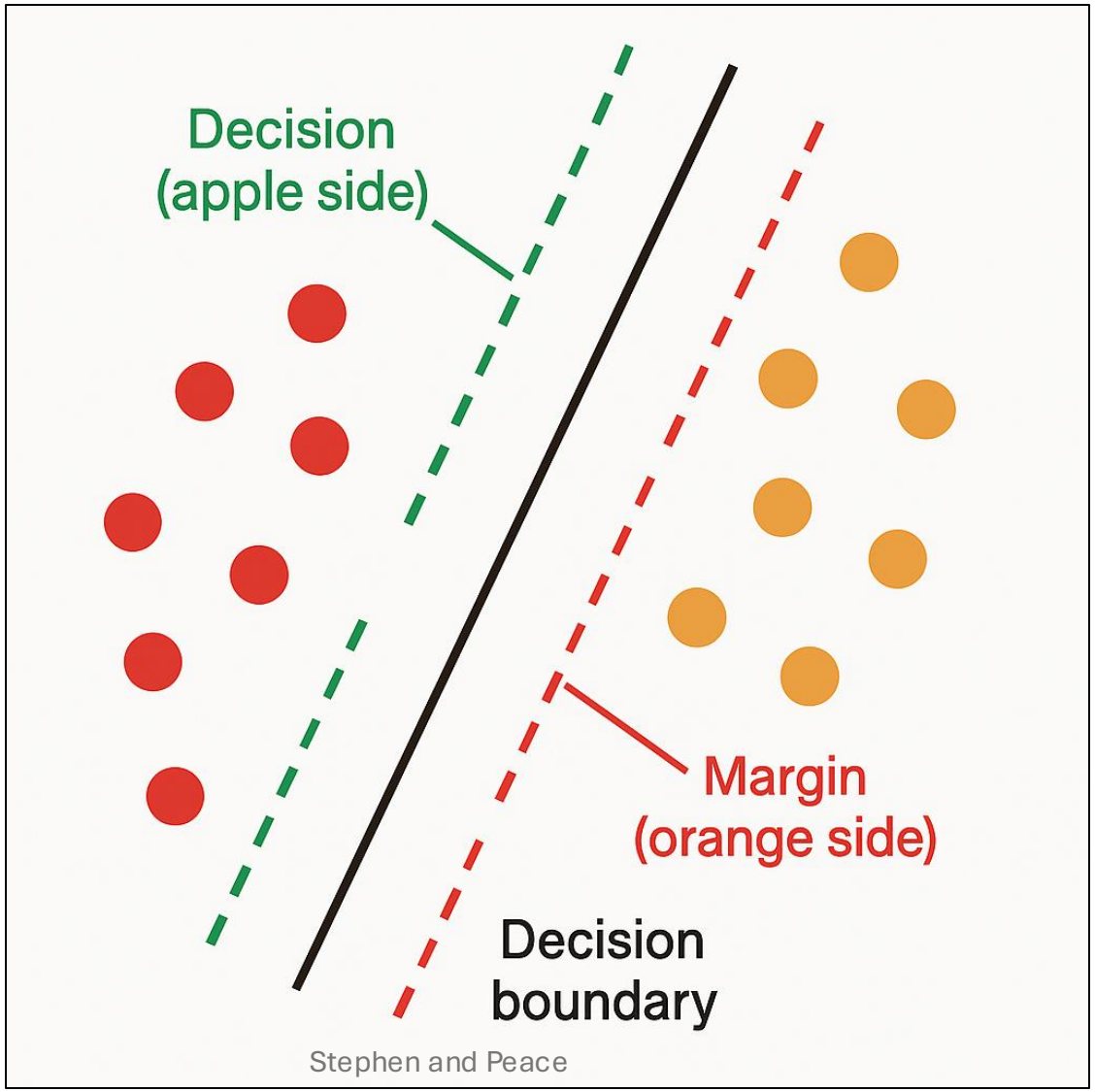
How it works?

- SVM imagines a boundary that's far from both groups.

- The closest points (one from each group) are called "support vectors" , they are like referees marking the edge of the field.

- The SVM positions the line so it's **as far from those referees as possible**.

Why “Best” Line?



Stephen and Peace

SVM algorithm

- **Measure two things** about each fruit — for example, weight and color shade.
- **Plot them on a chart:** apples cluster in one area, oranges in another.
- **SVM's job**: Find the best **straight line** (or curve) that divides apples from oranges.
- **Maximize the gap** between this line and the nearest fruit from each side — this gap is called the **margin**.
- **Support vectors** are the fruits that sit right at the edge of that margin — they "support" the line's position.
- When a **new fruit** comes in, SVM checks which side of the line it falls on to decide if it's an apple or an orange.

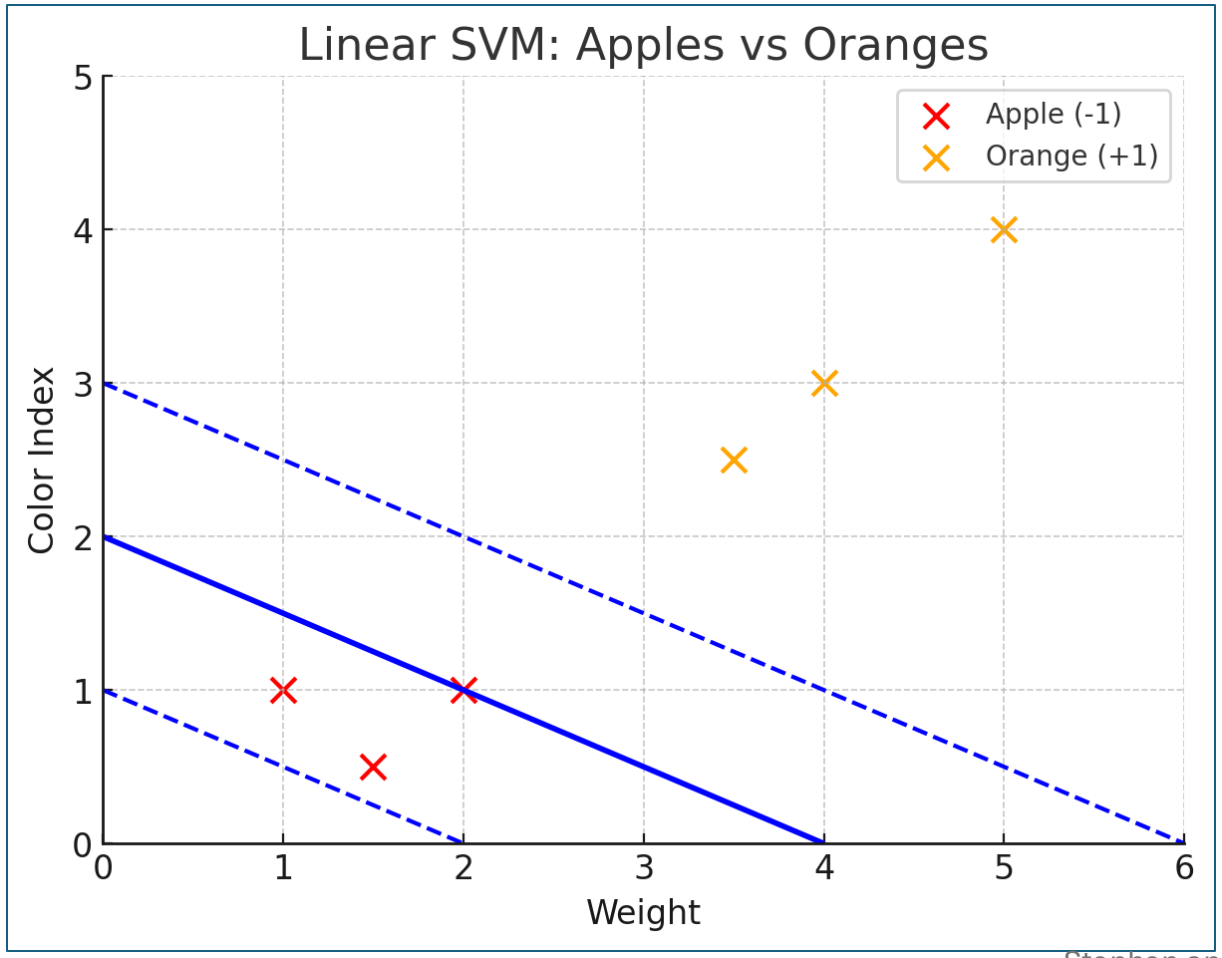
SVM classification for Orange vs Apple

- **SVM decision rule**
- In a **linear SVM**, we classify a fruit x as:

$$\hat{y} = sign(w \cdot x + b)$$

- w·x = **dot product** of the weight vector w and the fruit's features x
- b = bias term (shifts the boundary)

SVM classification for Orange vs Apple



- **Red dots** = Apples (class −1)
- **Orange dots** = Oranges (class +1)
- **Blue solid line** = Decision boundary $(w \cdot x + b = 0)$
- **Blue dashed lines** = Margins $(w \cdot x + b = \pm 1)$
- Points above the solid line are classified as **Orange**, and those below as **Apple**, based on the sign of the dot product $w \cdot x + b$.

The separating line

- The boundary between Apples and Oranges is: $w \cdot x + b = 0$
- Points where: $w \cdot x + b > 0$ → **Orange (+1)**
- Points where: $w \cdot x + b < 0$ → **Apple (-1)**

Farm examples?

- *Identify diseased vs. healthy plants from leaf measurements.*
- *Separate fertile vs. less fertile plots for planting.*
- *Classify soil samples into “needs fertilizer” vs. “good to go.”*
- *Sort harvest quality into “export grade” vs. “local market.”*