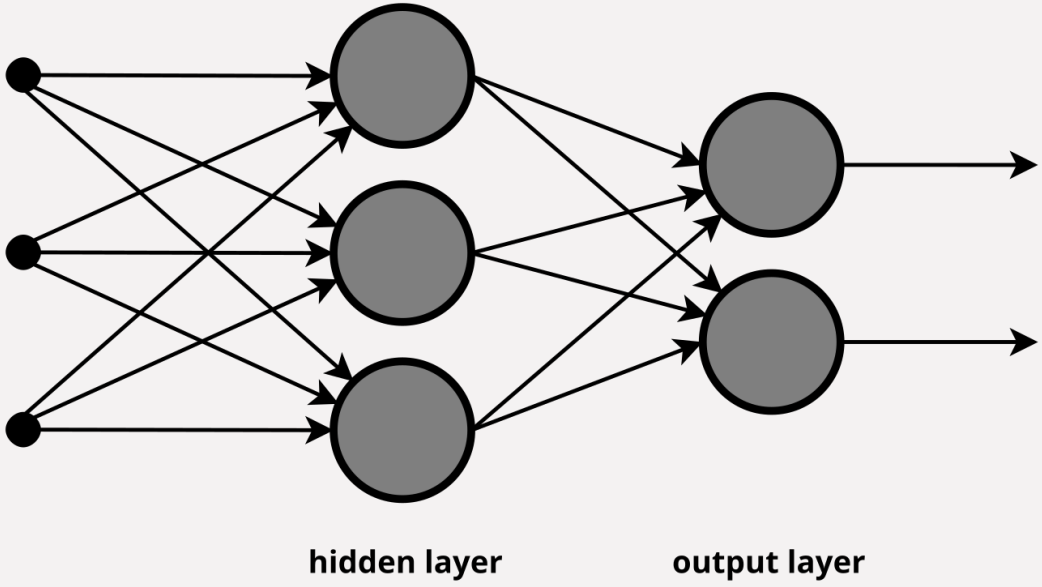
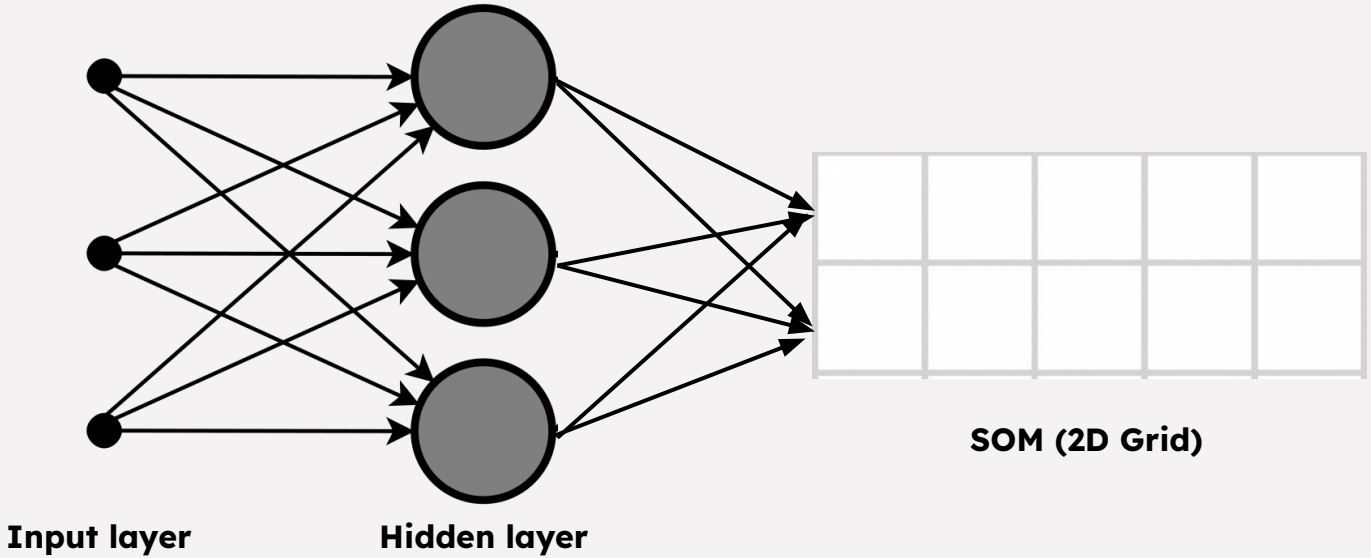
Self-Organizing Maps (SOM)

- Unsupervised learning
- Clustering
- Visualization of high-dimensional data

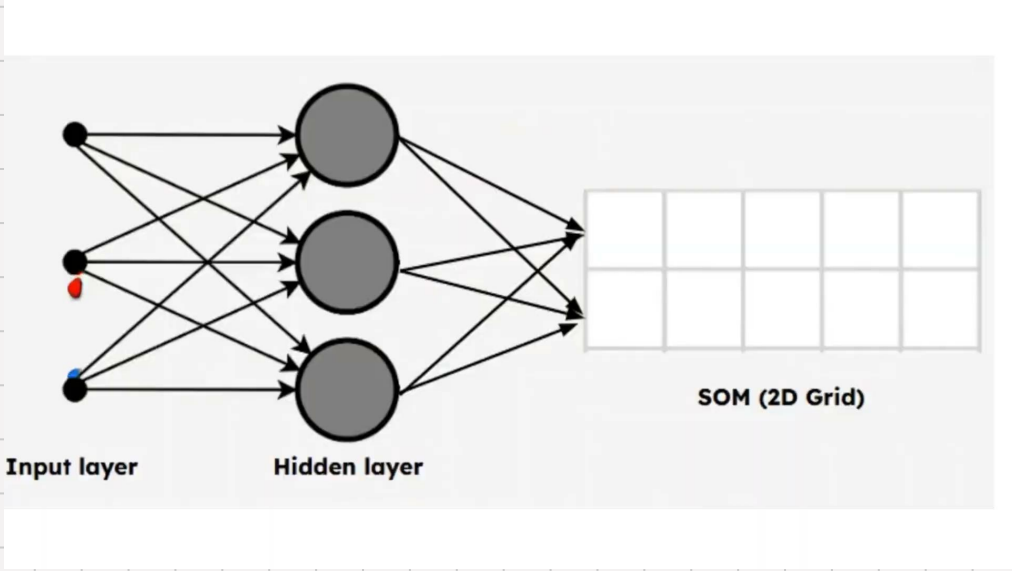


- Proposed by Teuvo Kohonen
- visualize complex data in a simpler 2D form
- Unsupervised Learning
- competitive learning



How SOM Works

1. Each neuron starts with random weights.
2. When a data point enters, the neuron most similar to it is selected.
3. The BMU and its neighboring neurons adjust their weights to become more like that data point.
4. Over time, the map organizes itself so that similar data points are close together.



SOM Training Algorithm

1. Initialization: Choose random values for the initial weights wij.
2. Sampling: Take a sample training input vector x(t) from the input layer.
3. Matching: Find the winning neuron from the output layer that has the weight vector closest to the input vector. It can be calculated by taking the square of the Euclidean distance for each output unit and finding the output unit that has the minimum Euclidean distance from the input vector.

   $$D_j = \Sigma(x_i - w_{ij}) \text{ for each } j = 1 \text{ to } m$$

4. New Weight Calculation: Find the new weights between the input vector sample and the winning output unit (neuron).

   $$w_{ij}(new) = w_{ij}(old) + n(t)(x_i - W_{ij}(old))$$

   where $n(t)$ is the learning rate.
5. Continuation: Repeat steps 2 to 4 until weight updation is negligible

Real-world Application

Customer Segmentation dataset:

Customer	Age	Annual Income	Spending Score
A	23	15k	80
B	42	70k	40
C	35	30k	60

1. Customer Segmentation
2. Disease Pattern Discovery
3. Image Compression & Recognition

Key points

1. SOM is an unsupervised learning algorithm
2. It reduces complex, high-dimensional data into a simple 2D map
3. Similar data points end up close together on the map
4. Discover hidden patterns