The hidden layer allows the network to:

* Create multiple decision boundaries

* Learn new features that are not present in the original input

* Combine them to form a non-linear boundary

Ex:-

→ Is x_1 OR x_2 active?

→ Is x_1 AND x_2 active?

Individually these are simple decisions

But together they capture the XOR structure