Model	Multi-Layer Perceptron (MLP)	Kolmogorov-Arnold Network (KAN)
Theorem	Universal Approximation Theorem	Kolmogorov-Arnold Representation Theorem
Formula (Shallow)	$f(\mathbf{x}) \approx \sum_{i=1}^{N(\epsilon)} a_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i)$	$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right)$
Model (Shallow)	fixed activation functions on nodes  learnable weights on edges	(b)  learnable activation functions on edges  sum operation on nodes
Formula (Deep)	$MLP(\mathbf{x}) = (\mathbf{W}_3 \circ \sigma_2 \circ \mathbf{W}_2 \circ \sigma_1 \circ \mathbf{W}_1)(\mathbf{x})$	$KAN(\mathbf{x}) = (\mathbf{\Phi}_3 \circ \mathbf{\Phi}_2 \circ \mathbf{\Phi}_1)(\mathbf{x})$
Model (Deep)	(c) $W_3$ $\sigma_2$ nonlinear, fixed $W_2$ $\sigma_1$ linear, learnable $\sigma_2$	(d) $\Phi_{3} = \begin{pmatrix} \mathbf{KAN}(\mathbf{x}) \\ \mathbf{\Phi}_{2} \end{pmatrix} \qquad \begin{pmatrix} \mathbf{nonlinear}, \\ \mathbf{learnable} \\ \mathbf{X} \end{pmatrix}$