Università degli Studi di Napoli Federico II

Dipartimento di Scienze Economiche e Statistiche

Corso di laurea Magistrale in Economia e Commercio

Metodi Statistici per il Data Mining

KOLMOGOROV-ARNOLD NETWORK VS. MLP: ANALISI DI UN NUOVO MODELLO DI RETI NEURALI

Relatore
Prof.ssa Rosaria Romano

Candidato
Luca Ruocco Matr.N28002171

Anno Accademico 2023\2024

Obiettivi della Tesi

Kolmogorov Arnold Network VS Multi Layers Perceptron

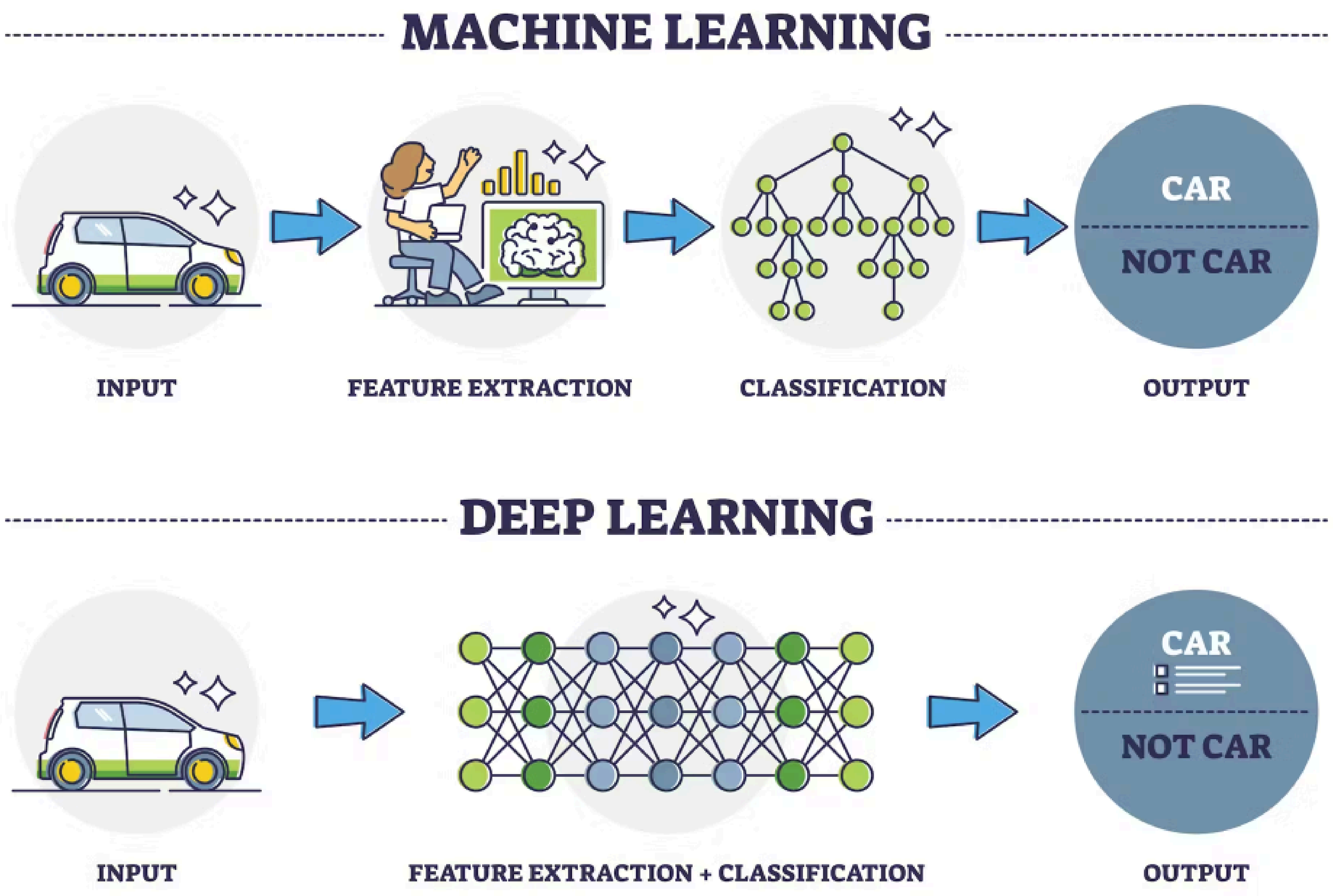
Accuratezza : percentuale di previsioni corrette.

Efficienza Computazionale: eseguire i calcoli rapidamente e con un uso minimo di risorse.

Interpretabilità: comprendere e spiegare il funzionamento dei modello

Il Deep Learning

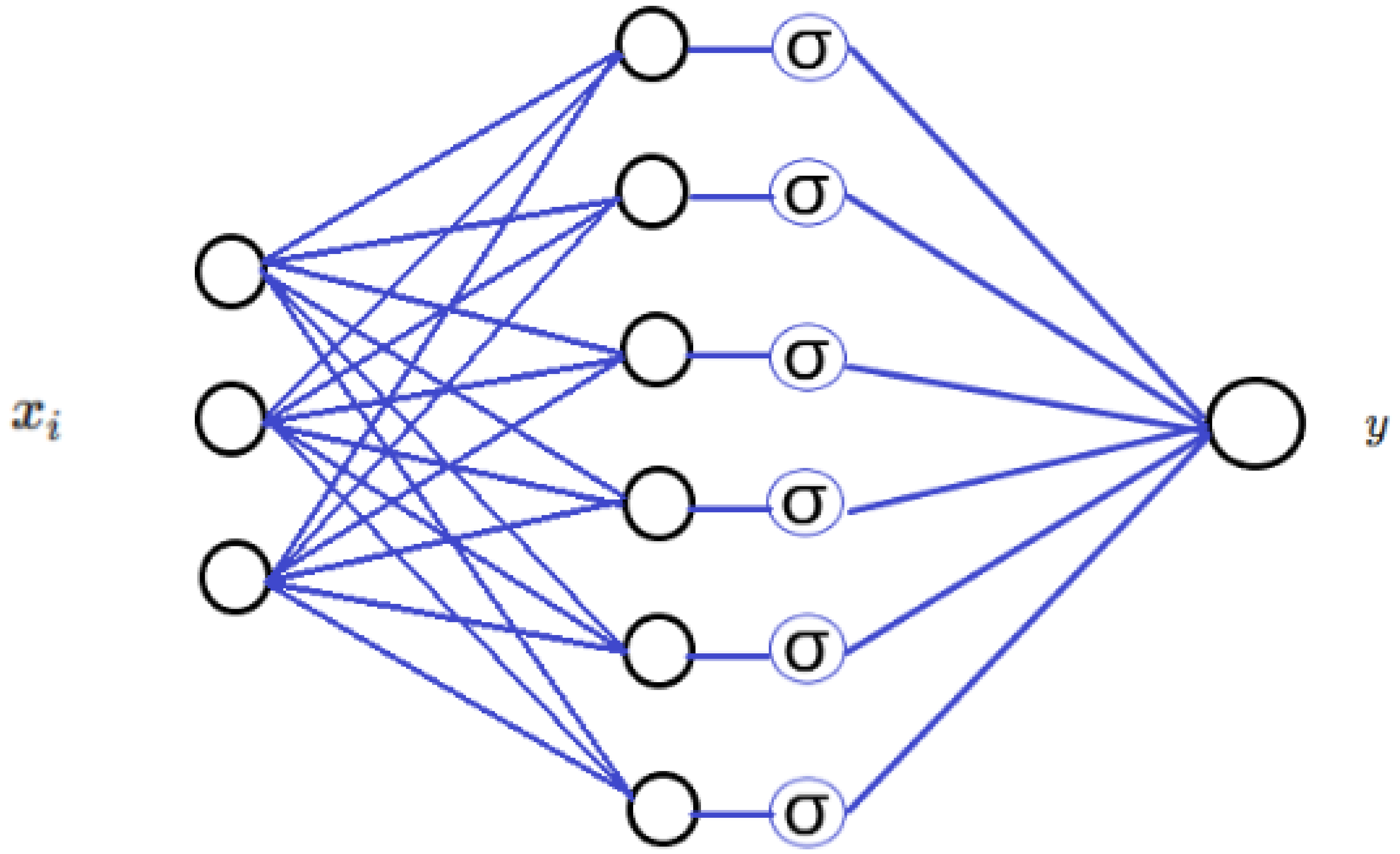
Il Deep Learning è una branca del Machine Learning che si concentra sull'uso di reti neurali profonde.
Si distingue per la capacità di apprendere autonomamente rappresentazioni complesse dei dati, senza la necessità di una significativa pre-elaborazione manuale delle feature.



Le Reti Neurali

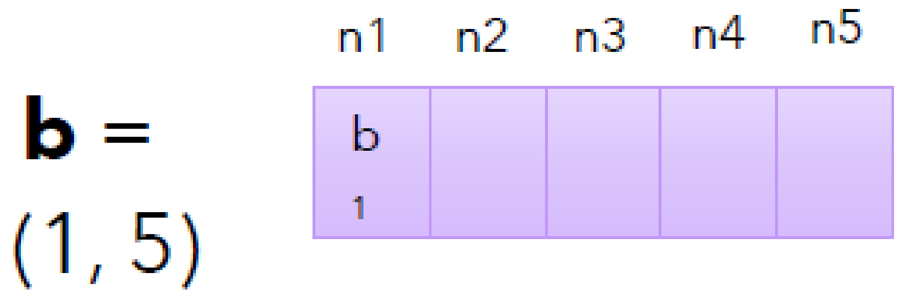
- Fondamento Teorico : Teorema Approssimazione Universale

- Come si presenta una rete neurale:

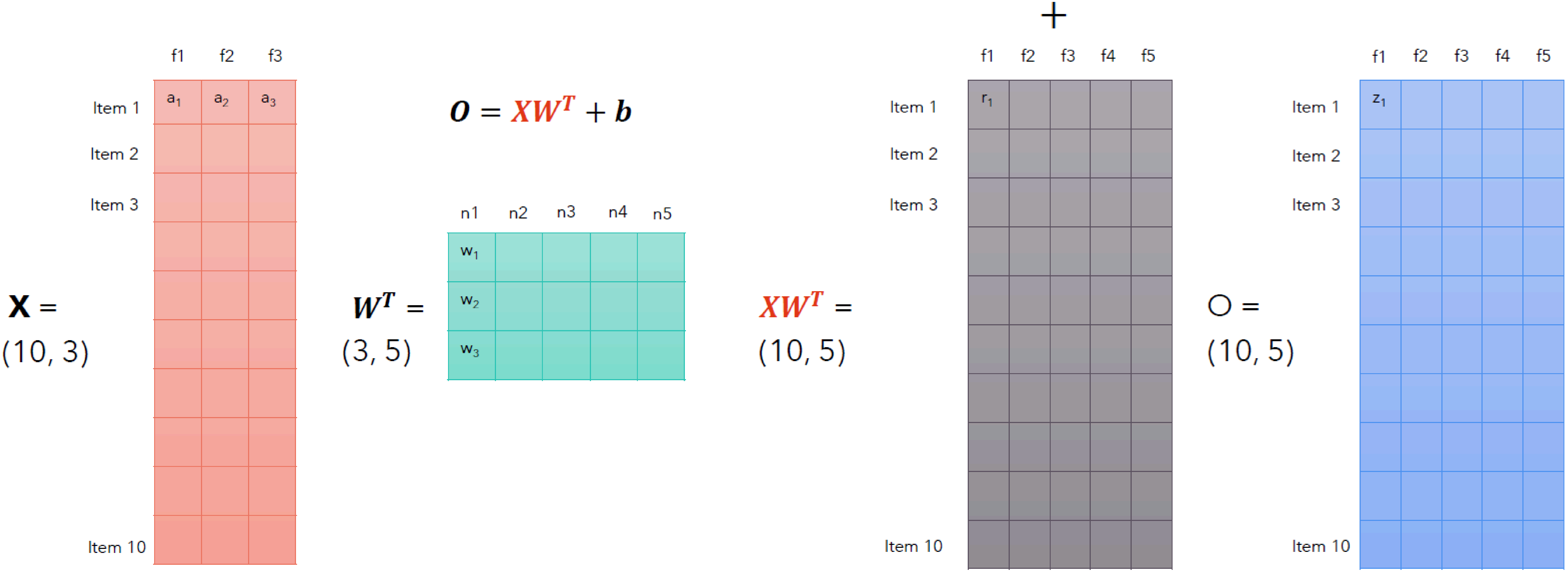


Funzionamento della Rete neurale

$$z_1 = (r_1 + b_1) = \left(\sum_{i=1}^{3} a_i w_i + b_1\right)$$

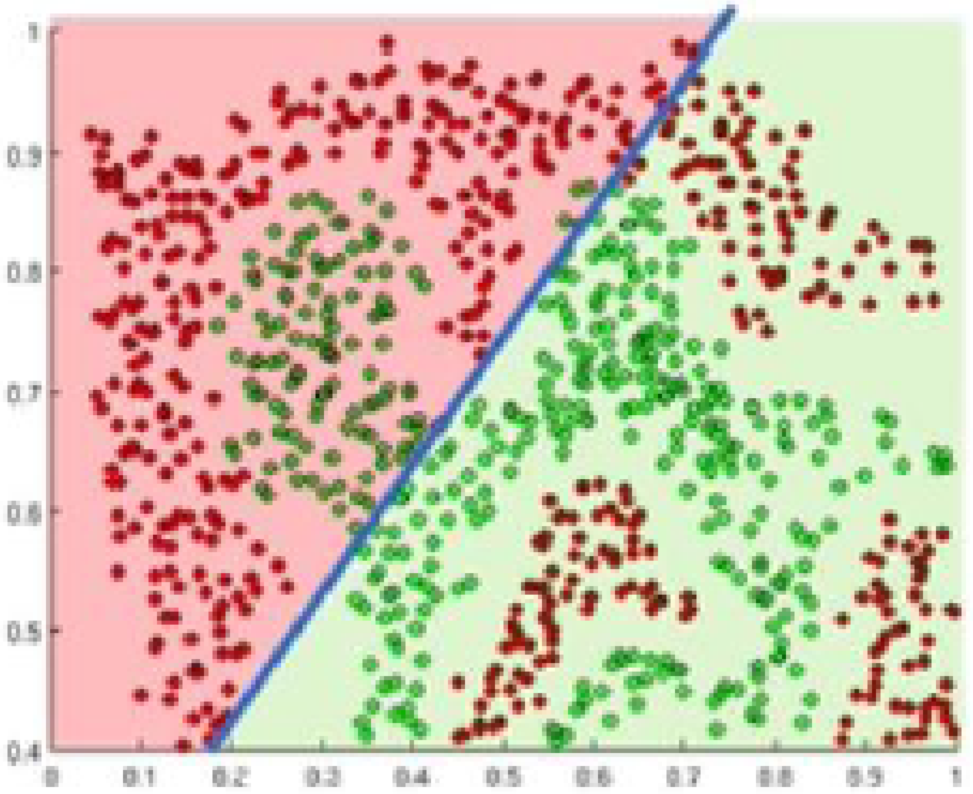


$$\boldsymbol{O} = \boldsymbol{XW^T} + \boldsymbol{b}$$



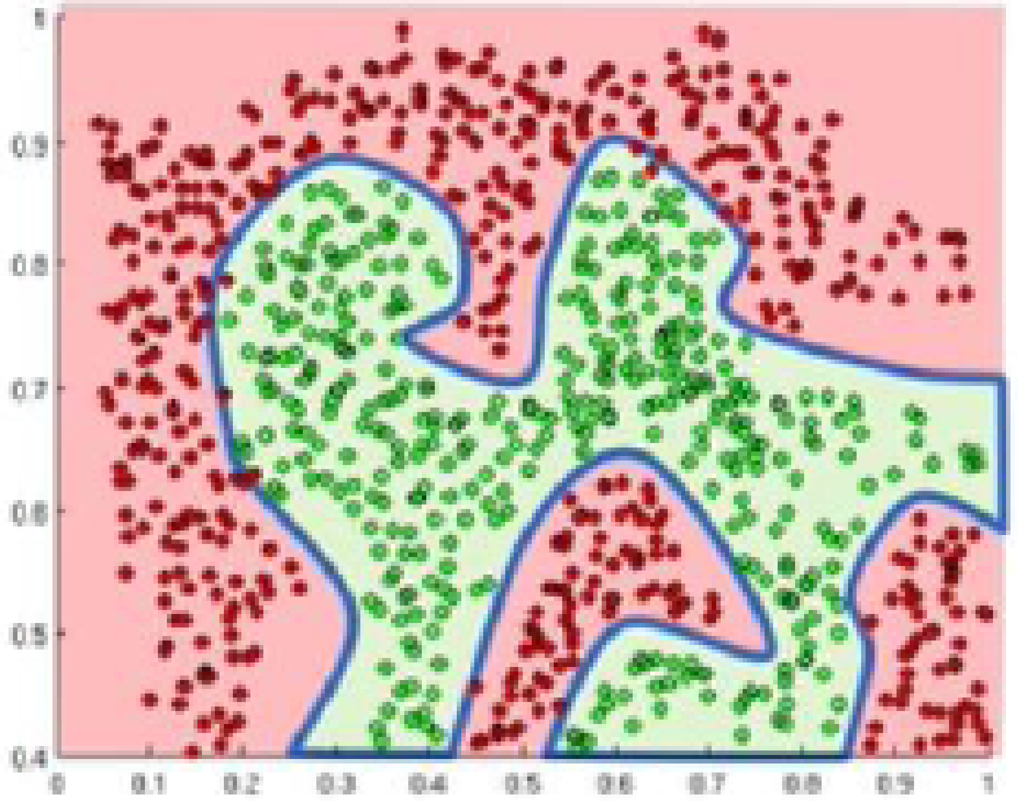
Importanza delle Funzioni di Attivazione

$$O_1 = xW_1^T + b_1$$



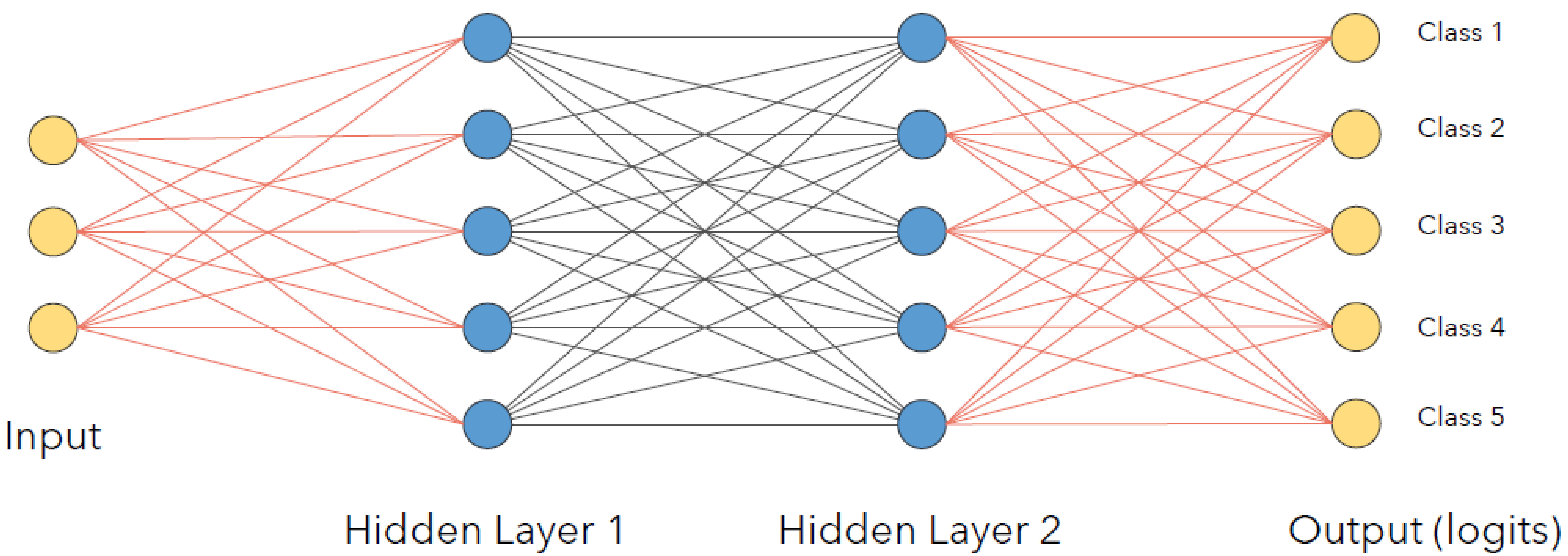
$$O_2 = (O_1)W_2^T + b_2$$

$$O_2 = (xW_1^T + b_1)W_2^T + b_2$$



$$O_2 = xW_1^TW_2^T + b_1W_2^T + b_2$$

Multi Layers Perceptron

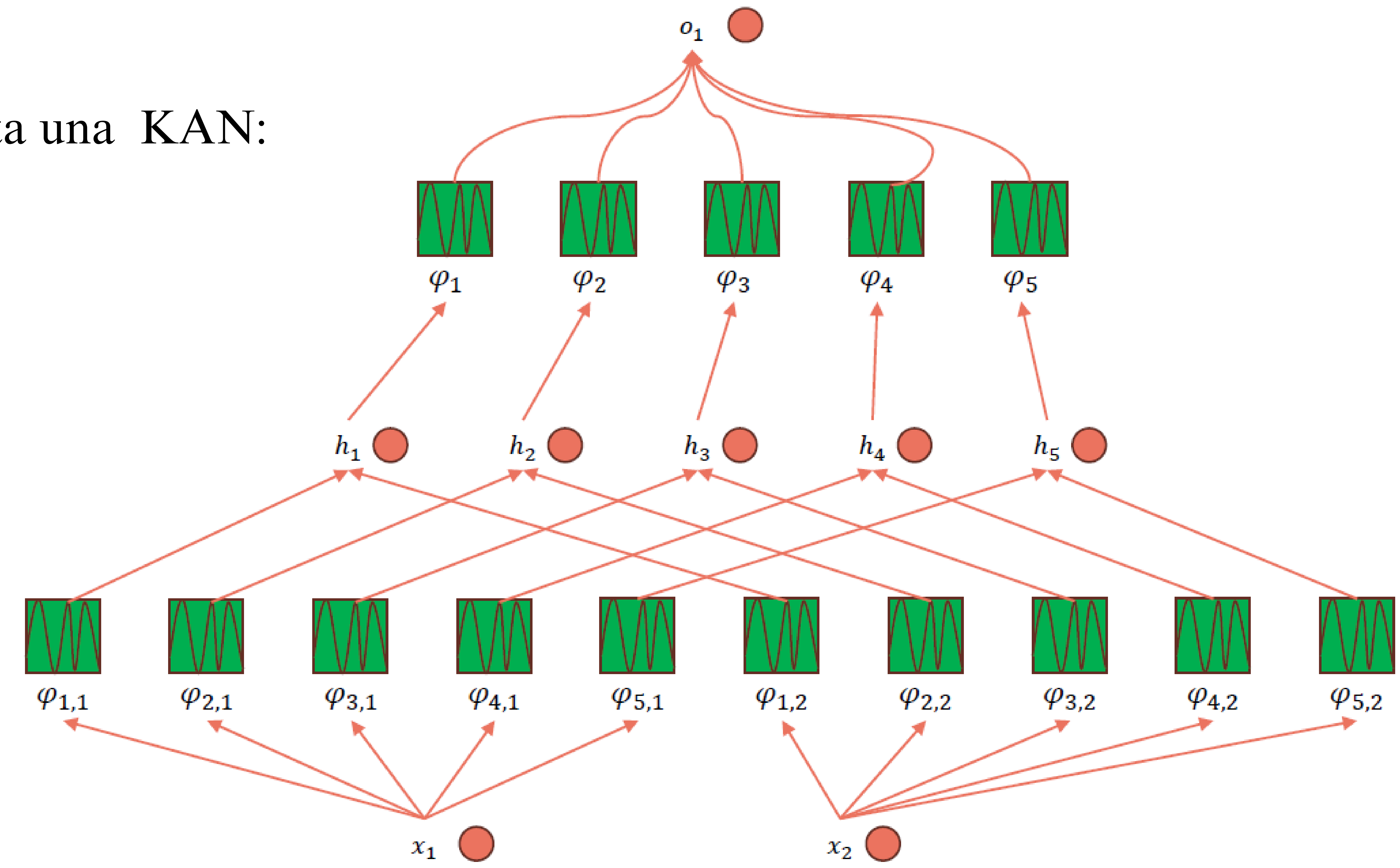


Vantaggi	Svantaggi
Versatile per risolvere problemi come classificazione, regressione	Scalabilità limitata
Capacità di apprendere relazioni non lineari complesse	Costo computazionale elevato con l'aumento di strati e neuroni
Generalizzazione efficace su dati non visti	Difficoltà nella selezione degli iperparametri
	Limiti nella gestione di strutture dati complesse o dinamiche

KOLMOGOROV-ARNOLD NETWORK (KAN)

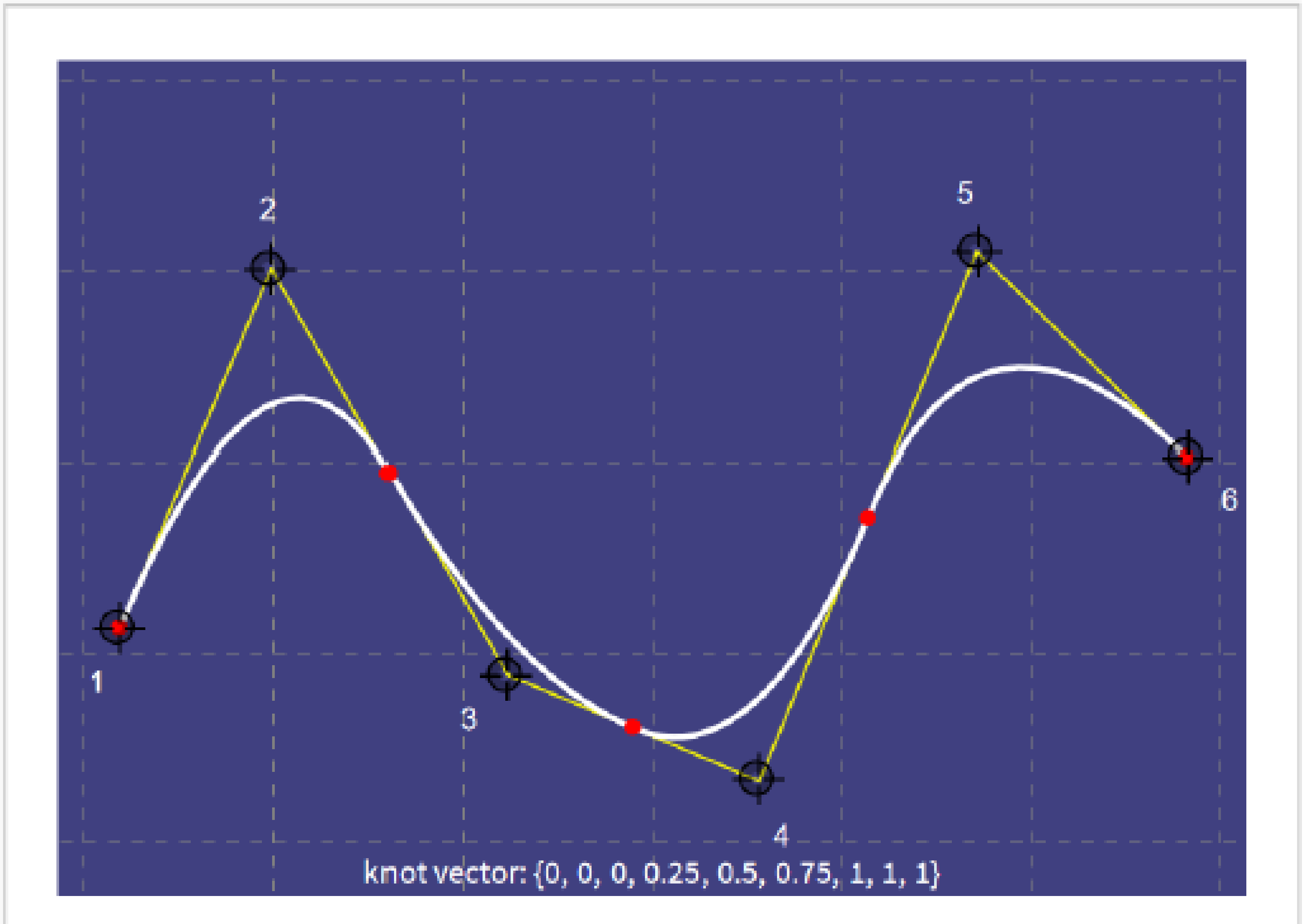
- Fondamento Teorico: Teorema di Rappresentazione di Kolmogorov - Arnold

Come si presenta una KAN:

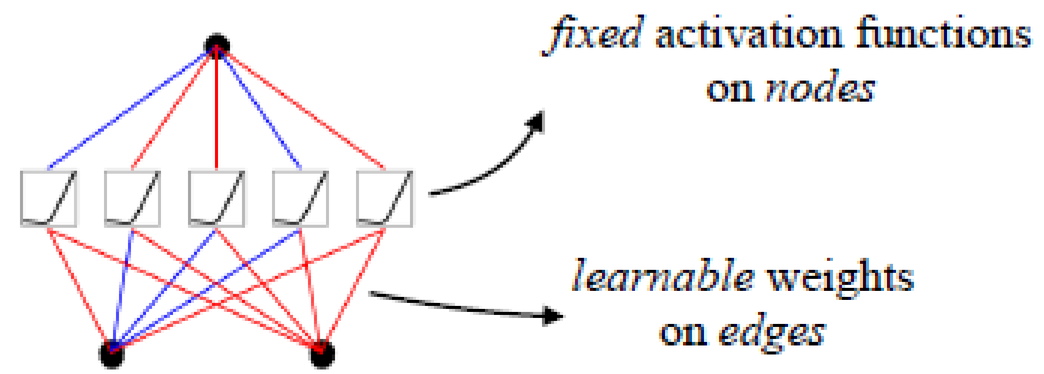
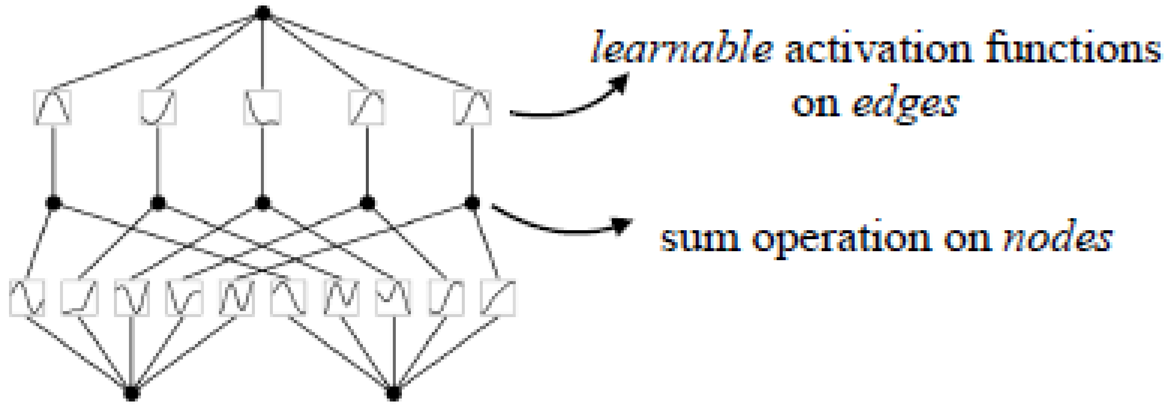
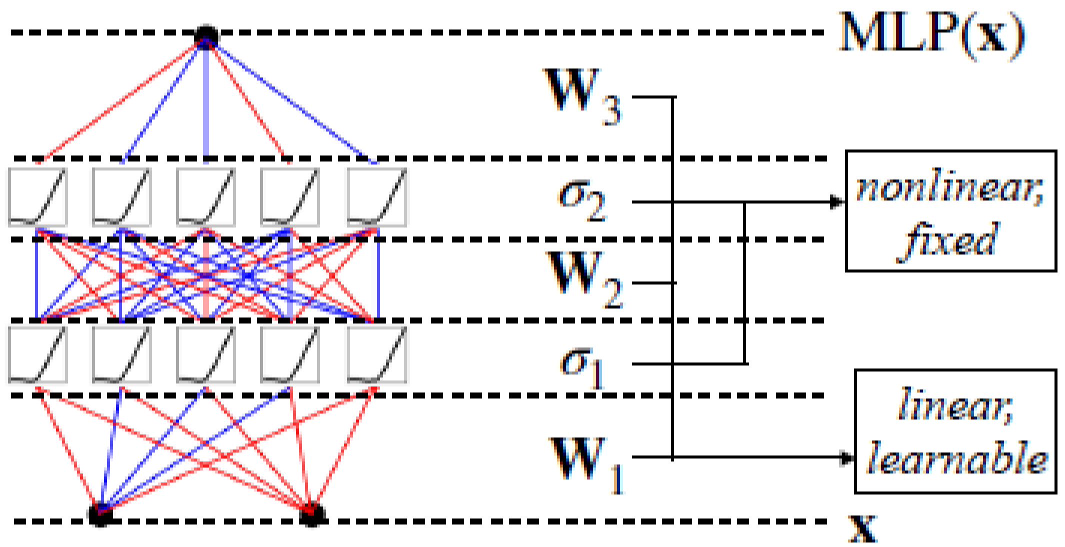
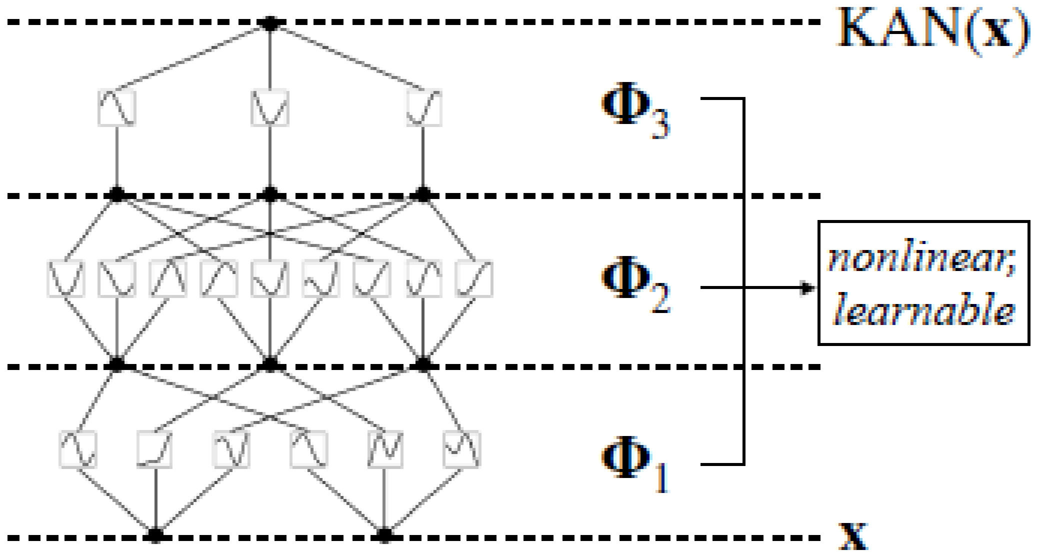


Importanza delle B-Spline

- Le B-spline sono un elemento essenziale nelle reti Kolmogorov-Arnold, in quanto vengono utilizzate per rappresentare le funzioni univariate che compongono il modello.



KAN vs MLP

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)	(a) 	(b) 
Formula (Deep)	$\text{MLP}(\mathbf{x}) = (\mathbf{W}_3 \circ \sigma_2 \circ \mathbf{W}_2 \circ \sigma_1 \circ \mathbf{W}_1)(\mathbf{x})$	$\text{KAN}(\mathbf{x}) = (\mathbf{\Phi}_3 \circ \mathbf{\Phi}_2 \circ \mathbf{\Phi}_1)(\mathbf{x})$
Model (Deep)	(c) 	(d) 

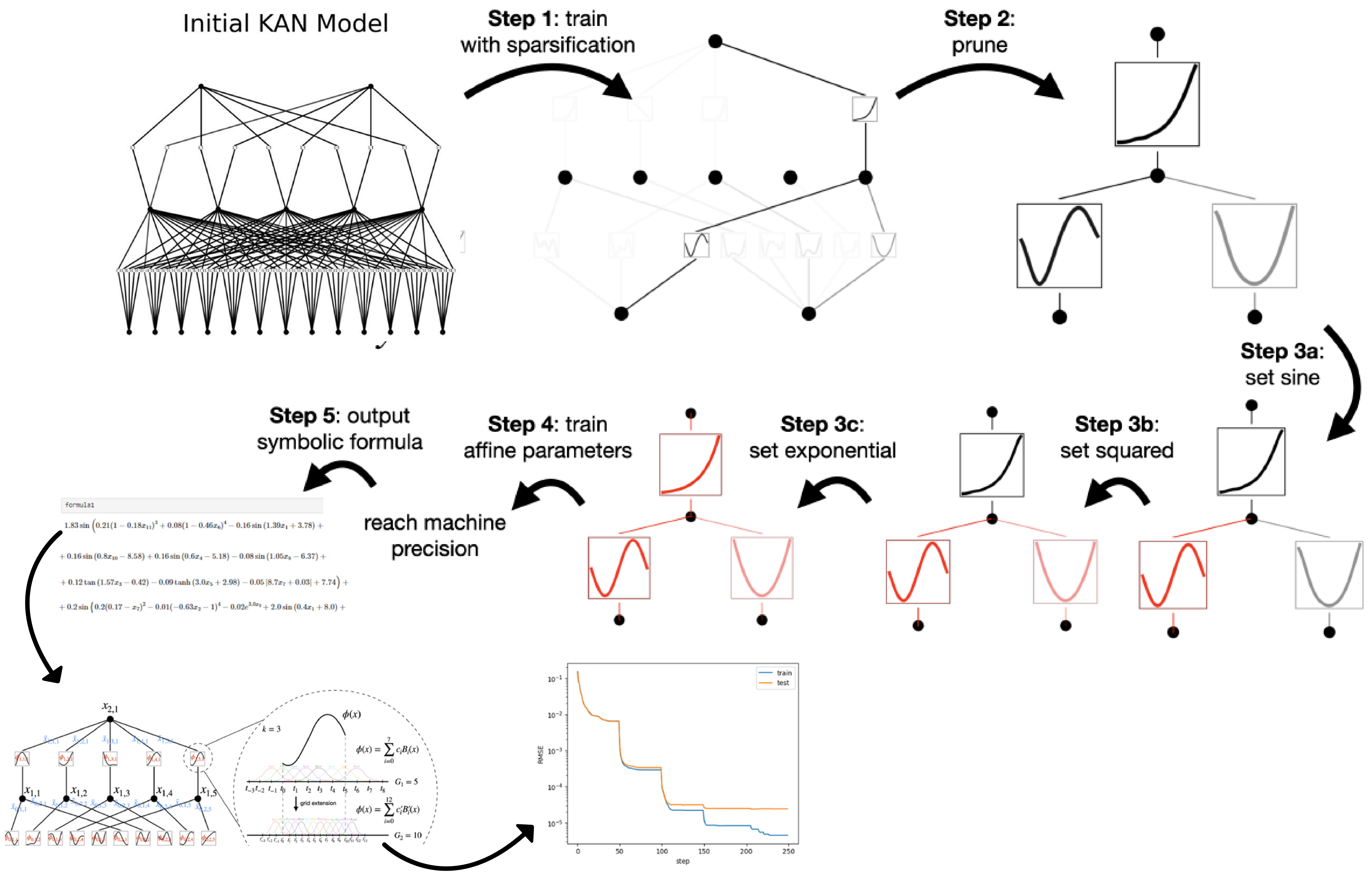
Dataset Utilizzati

Dataset	n	% classi	% var. numeriche	% var. categoriche
Heart Failure	5000	69% Survived 31% Failure	54%	46%
Breast Cancer	569	63% Benign 37% Malign	100%	0%
Iris	150	33% Setosa 33% Versicolor 33%Virginica	100%	0%
Bank Marketing	4521	88% no 12% yes	37.50%	62.50%
Heart Disease	1190	47% Normal 53% Heart Disease	41.67%	58.33%
Fetal Health	2126	78% Normal 14% Suspicious 8% Pathological	100%	0%
Cervical Cancer	858	94% Healthy 6% Cancer	61.54%	38.46%

kaggle

UCI Machine Learning Repository

Heart Disease

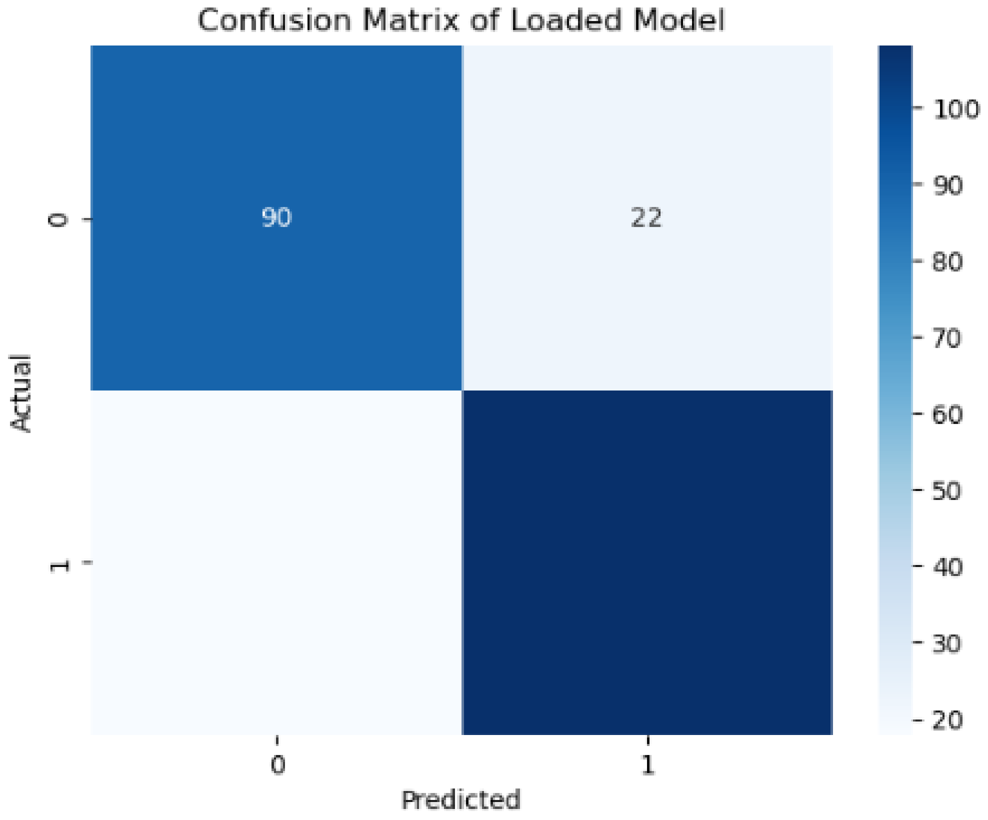


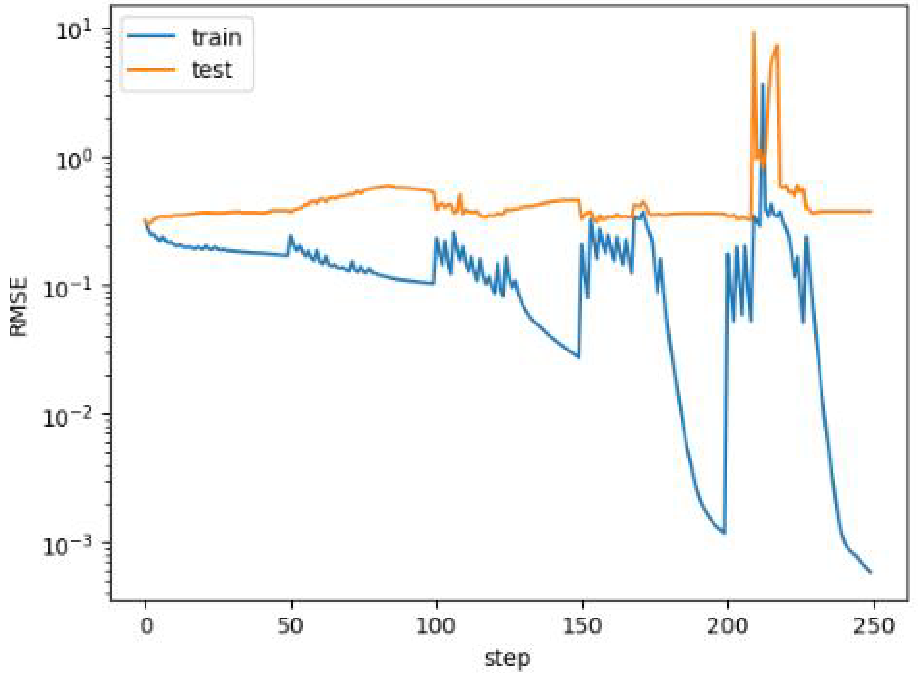
Heart Disease 2

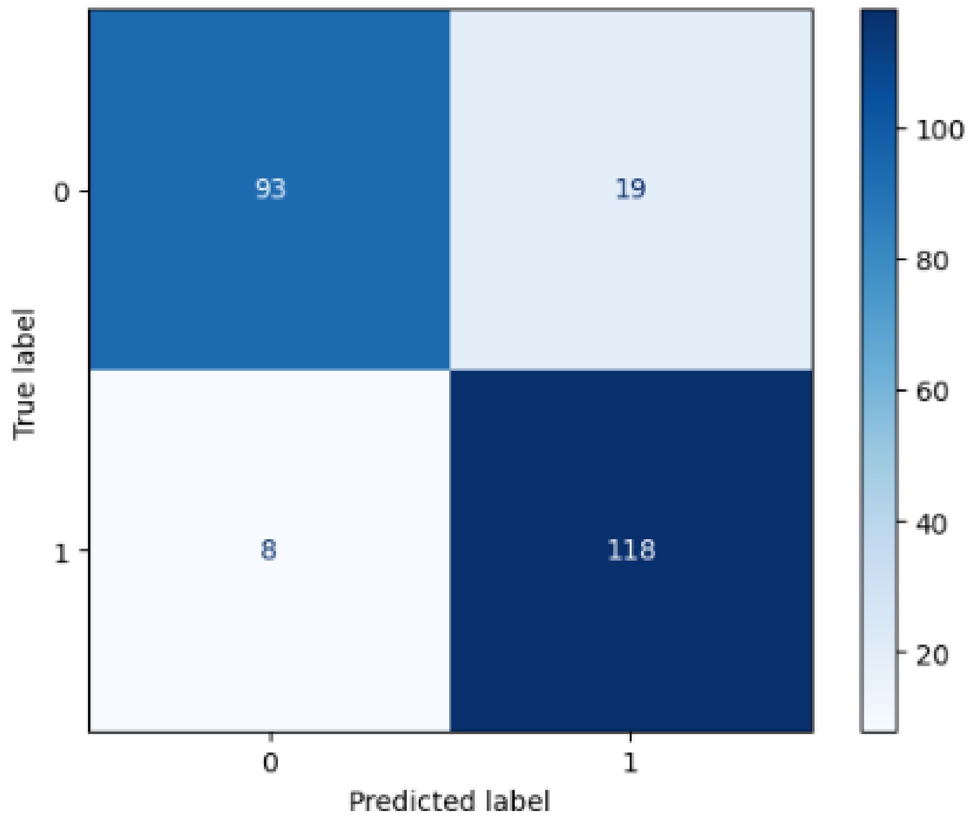
MLP

KAN

```
Confusion Matrix:
[[ 90  22]
 [ 18 108]]
```







Risultati dei due modelli

Dataset	Errori Classi (Test)	Accuratezza (Training)	Accuratezza (Test)
Heart Failure	6,56% Normal - 38,36% Failure	81,50%	79,45%
Breast Cancer	1,38% Benign - 13,28% Malign	97,32%	93,02%
Iris	0 %Setosa - 70% Verticolor - 0% Virginica	86,12%	76,92%
Bank Marketing	5,74% No - 51,91% Yes	86,96%	72,49%
Heart Disease	19,64% Normal - 14,28% Heart Disease	83,19%	78,92%
Featal Health	4,82% Normal - 40,68% Suspicious - 76,2% Pathological	82,63%	77,13%
Cervical Cancer	0% Healthy - 100% Cancer	93,60%	85,22%

MLP

KAN

Dataset	Errori Classi (Test)	Accuratezza (Training)	Accuratezza (Test)
Heart Failure	3,35% Normal - 6,68% Failure	96,97%	95,60%
Breast Cancer	0% Benign - 9,52% Malign	99,12%	96,49%
Iris	0 %Setosa - 5% Verticolor - 5% Virginica	100%	93,33%
Bank Marketing	4,24% No - 30,76% Yes	91,15%	88,29%
Heart Disease	16,96% Normal - 6,34% Heart Disease	93,17%	88,66%
Featal Health	5,42% Normal - 25,4% Suspicious - 31,43% Pathological	93,12%	89,67%
Cervical Cancer	0% - Healty - 92,3% Cancer	94,33%	87,46%

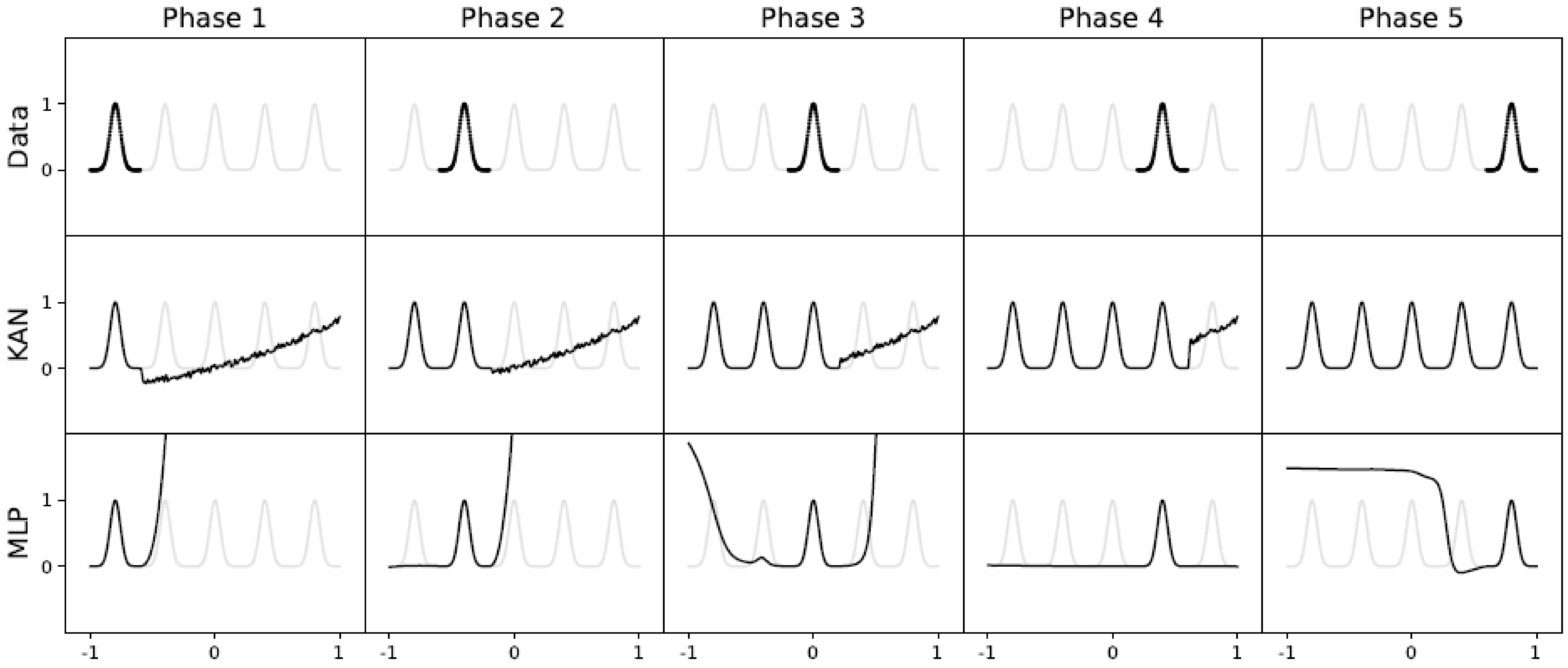
Conclusioni

Accuratezza : Le reti KAN ottengono risultati molto superiori e sono in grado di generalizzare meglio

Efficienza Computazionale: A parità di condizioni le MLP risultano meno costose in termini computazionali

Interpretabilità: Seppur limitata la fase di pruning ed estrazione di formula simbolica delle KAN risulta intuitiva

Continual Learning



Università degli Studi di Napoli Federico II **Dipartimento di Scienze Economiche e Statistiche**

Corso di laurea Magistrale in Economia e Commercio

Metodi Statistici per il Data Mining

KOLMOGOROV-ARNOLD NETWORK VS. MLP: ANALISI DI UN NUOVO MODELLO DI RETI NEURALI

GRAZIE PER L'ATTENZIONE

Luca Ruocco

RIFERIMENTI

Agliari, E., Alessandrelli, A., Barra, A., Centonze, M. S., & Ricci-Tersenghi, F. (2024). Generalized hetero-associative neural networks. arXiv preprint arXiv:2409.08151.

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of microbiological methods, 43(1), 3-31.

Cilimkovic, M. (2015). Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin, 15(1).

Fine, T. L. (2006). Feedforward neural network methodology. Springer Science & Business Media.

Günther, F., & Fritsch, S. (2010). Neuralnet: training of neural networks. R J., 2(1), 30.

Kelleher, J. D. (2019). Deep learning. MIT press.

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. Artificial Neural Networks: Methodological Advances and Biomedical Applications. InTech, 1-18.

Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., ... & Tegmark, M. (2024). Kan: Kolmogorov-arnold networks. arXiv preprint arXiv:2404.19756.

Mercioni, M. A., & Holban, S. (2020, May). The most used activation functions: Classic versus current. In 2020 International Conference on Development and Application Systems (DAS) (pp. 141-145). IEEE.

Nishijima, T. (2021). Universal approximation theorem for neural networks. arXiv preprint arXiv:2102.10993.

Wang, H., & Raj, B. (2017). On the origin of deep learning. arXiv preprint arXiv:1702.07800.

Wu, Y. C., & Feng, J. W. (2018). Development and application of artificial neural network. Wireless Personal Communications, 102, 1645-1656.