🧠 CLASS 1: ADVANCED ACTIVATION FUNCTIONS PART 4

Activation functions play a crucial role in neural networks by introducing non-linear properties that allow the model to learn complex patterns in the data. This session will explore additional advanced activation functions beyond the commonly used Sigmoid, Tanh, and ReLU.

🔍 Advanced Activation Functions

These functions are essential for deeper networks that need to model more complex relationships within large or highly varied datasets.

Overview of Additional Functions:

Leaky ReLU: Improves upon the ReLU function by allowing a small, non-zero gradient when the unit is not active and preventing the dying ReLU problem.

Parametric ReLU (PReLU): This variation of Leaky ReLU has a coefficient for the negative part of the function that is learned during training, potentially improving learning capability.

Exponential Linear Unit (ELU): ELU speeds up learning and achieves higher classification accuracy by reducing the vanishing gradient problem with an exponential curve for negative values, unlike ReLU.

Softplus: A smooth approximation to the ReLU function, defined as $\log(1 + e^x)$, which creates a smooth curve and reduces the likelihood of dead neurons.

Swish: A self-gated activation function that has been found to work better than ReLU in some deep learning models. It is defined as $x * \text{sigmoid}(\beta x)$.

📊 Importance and Selection

Choosing the right activation function can significantly influence the performance of a neural network. Each function has its strengths and weaknesses, and the choice often depends on the specific requirements of the application and the nature of the problem being solved.

📝 Conclusion

This overview of advanced activation functions introduces more tools available to neural network designers. Understanding these options expands the toolkit for designing networks that can learn from complex, high-dimensional datasets.