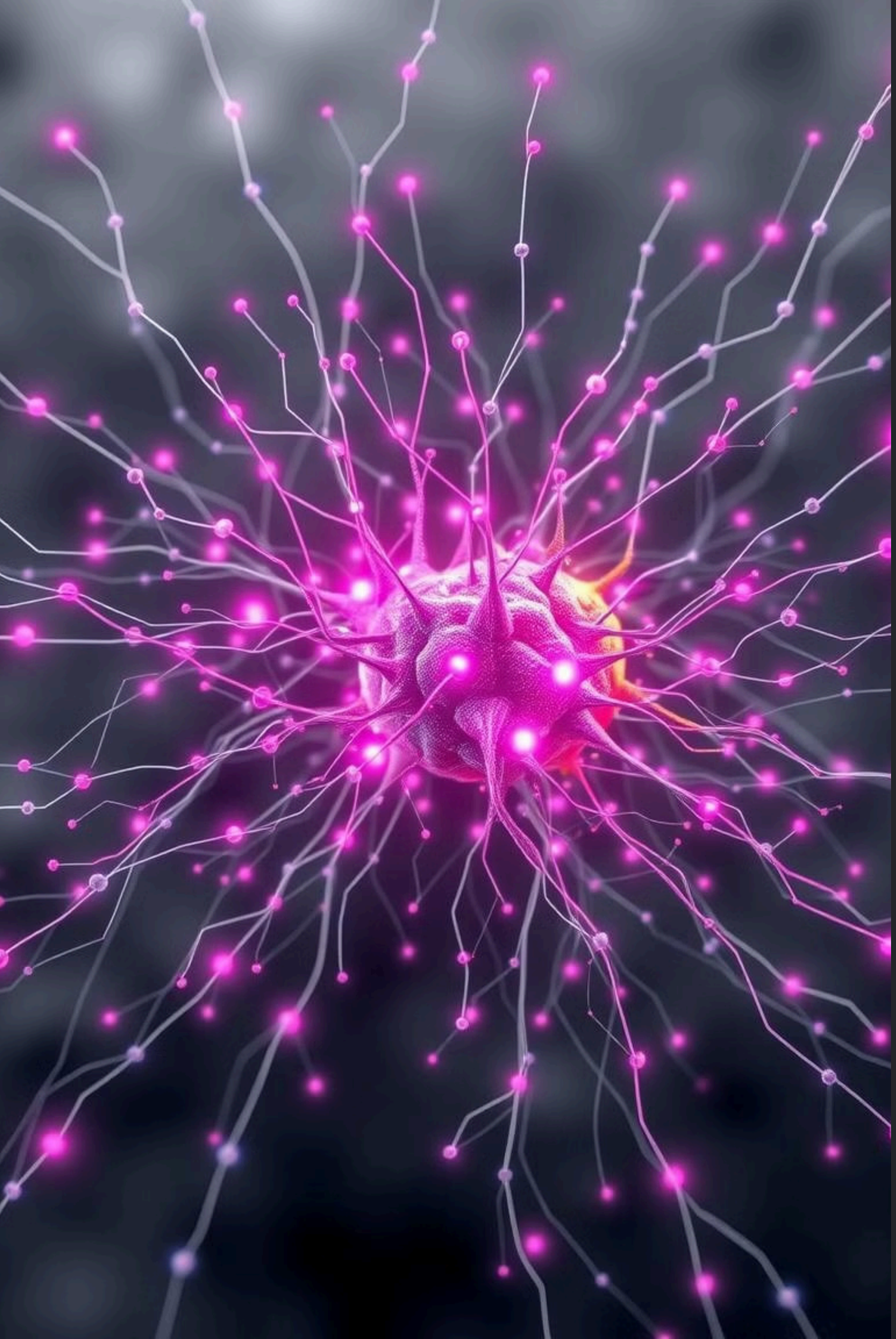


Activation Functions, Loss Functions, and Hyperparameters: A Deep Dive

Welcome! Today we'll explore the fundamental concepts of activation functions, loss functions, and hyperparameters in machine learning. These concepts are crucial for building effective and accurate models.

S **by Sumana Reddy**

What are Activation Functions and Why Do We Need Them?

Adding Non-Linearity

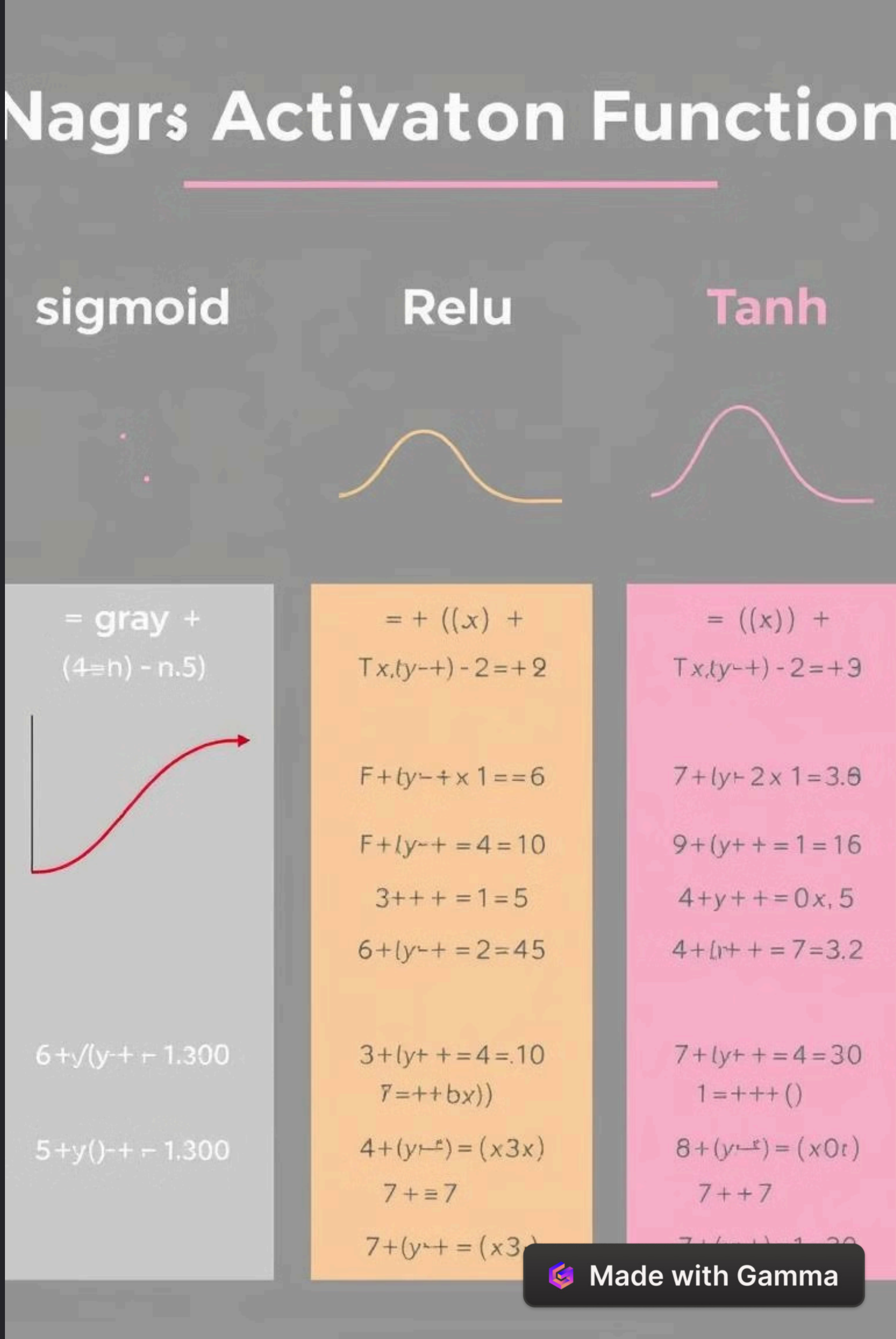
Activation functions introduce non-linearity into neural networks, allowing them to learn complex relationships in data. This is essential for modeling real-world phenomena.

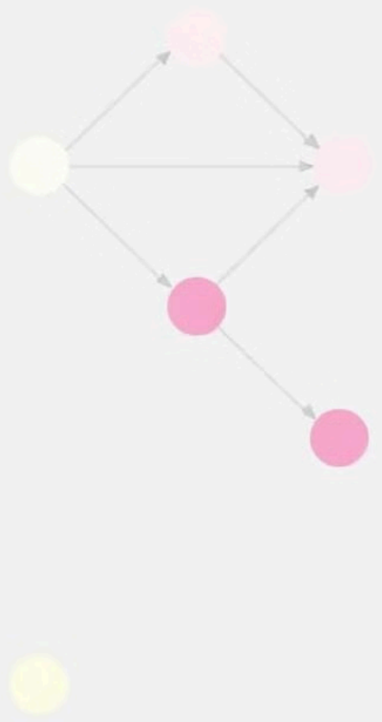
Decision Making

They enable neural networks to make decisions by transforming the output of neurons into probabilities or classifications.

Popular Activation Functions: Sigmoid, ReLU, and Beyond

1. Sigmoid: A classic function that produces a smooth, S-shaped curve. It's often used in output layers for classification tasks.
2. ReLU (Rectified Linear Unit): A simple function that outputs the input if it's positive, otherwise it outputs zero. It's widely used due to its efficiency and ability to prevent vanishing gradients.
3. Tanh (Hyperbolic Tangent): Similar to Sigmoid, but with a wider range of outputs. It can be more effective in certain scenarios, like recurrent neural networks.





Understanding Loss Functions: Measuring Model Performance

Quantifying Error

Loss functions quantify the error made by a machine learning model, comparing its predictions to actual values.

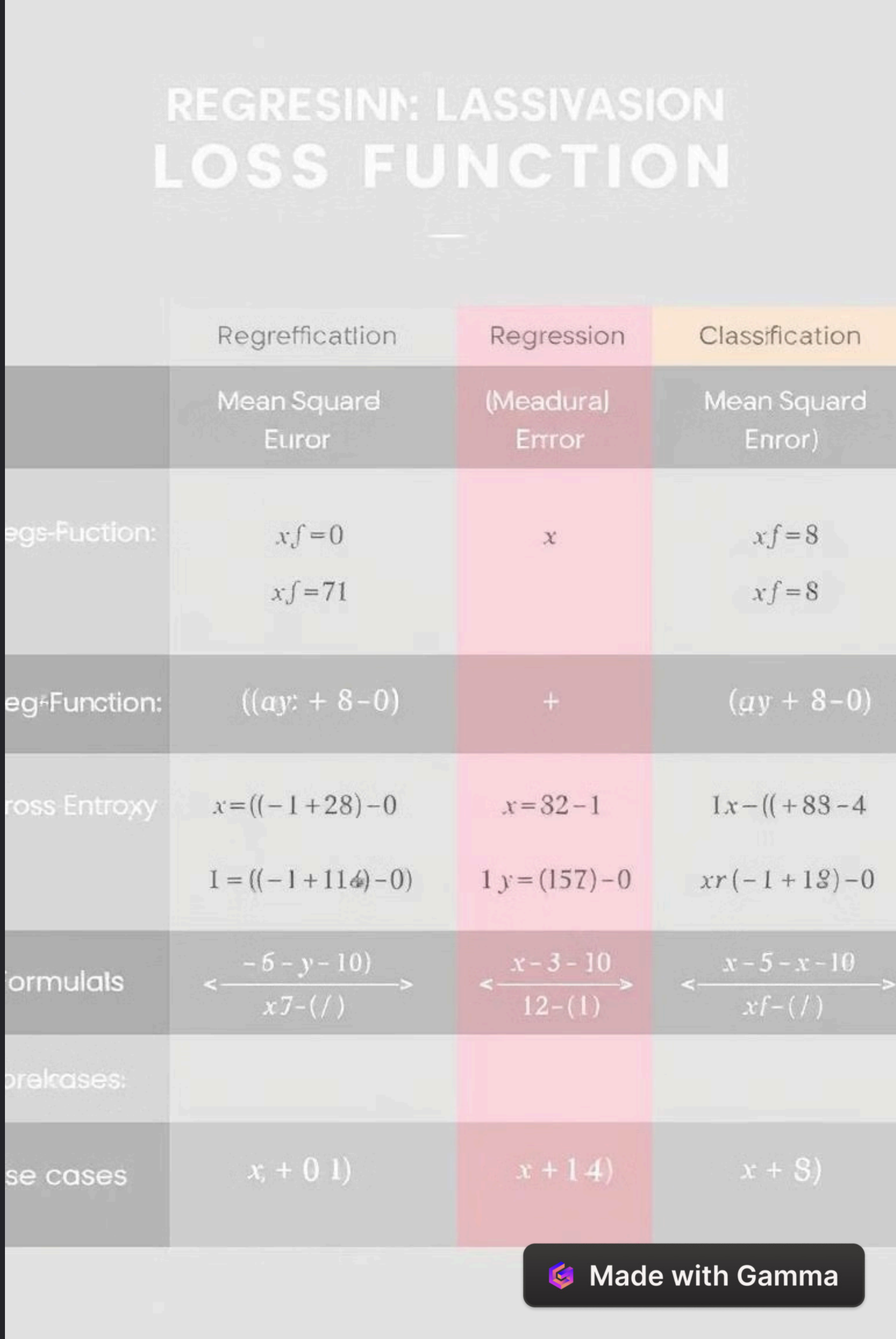
Optimizing Performance

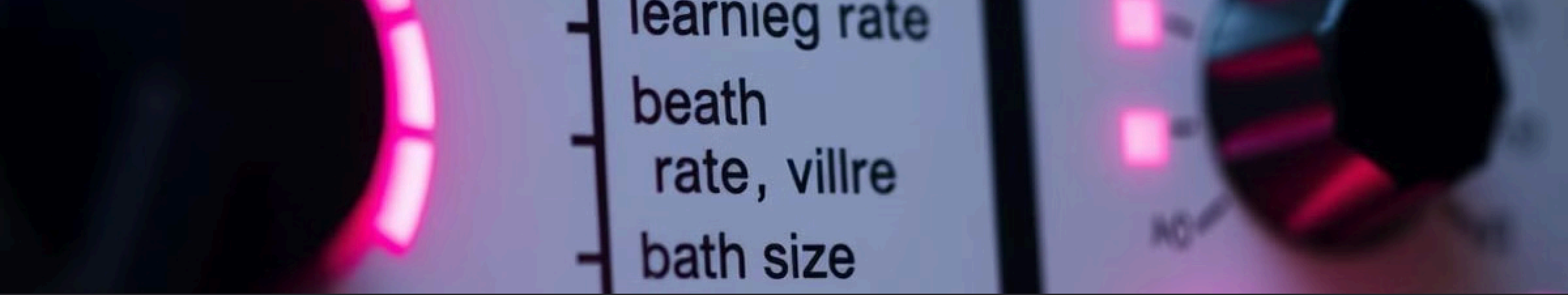
The goal of training a model is to minimize its loss function, improving its accuracy and overall performance.

Common Loss Functions: Regression vs. Classification

Regression Loss: Used in regression problems, where the goal is to predict continuous values. Common examples include Mean Squared Error (MSE) and Mean Absolute Error (MAE).

Classification Loss: Used in classification problems, where the goal is to predict discrete categories. Common examples include Cross-Entropy and Hinge Loss.





Hyperparameters: The Knobs You Can Tune

1 **Learning Rate**

Determines the step size used during optimization, controlling how quickly the model learns.

2 **Batch Size**

The number of samples used in each training iteration, impacting the training speed and generalization ability.

3 **Number of Layers**

The depth of the neural network, impacting its capacity to learn complex patterns.

Hyperparameter Tuning Techniques: Grid Search, Random Search, and Bayesian Optimization

1

Grid Search

A systematic approach that explores all possible combinations of hyperparameter values.

2

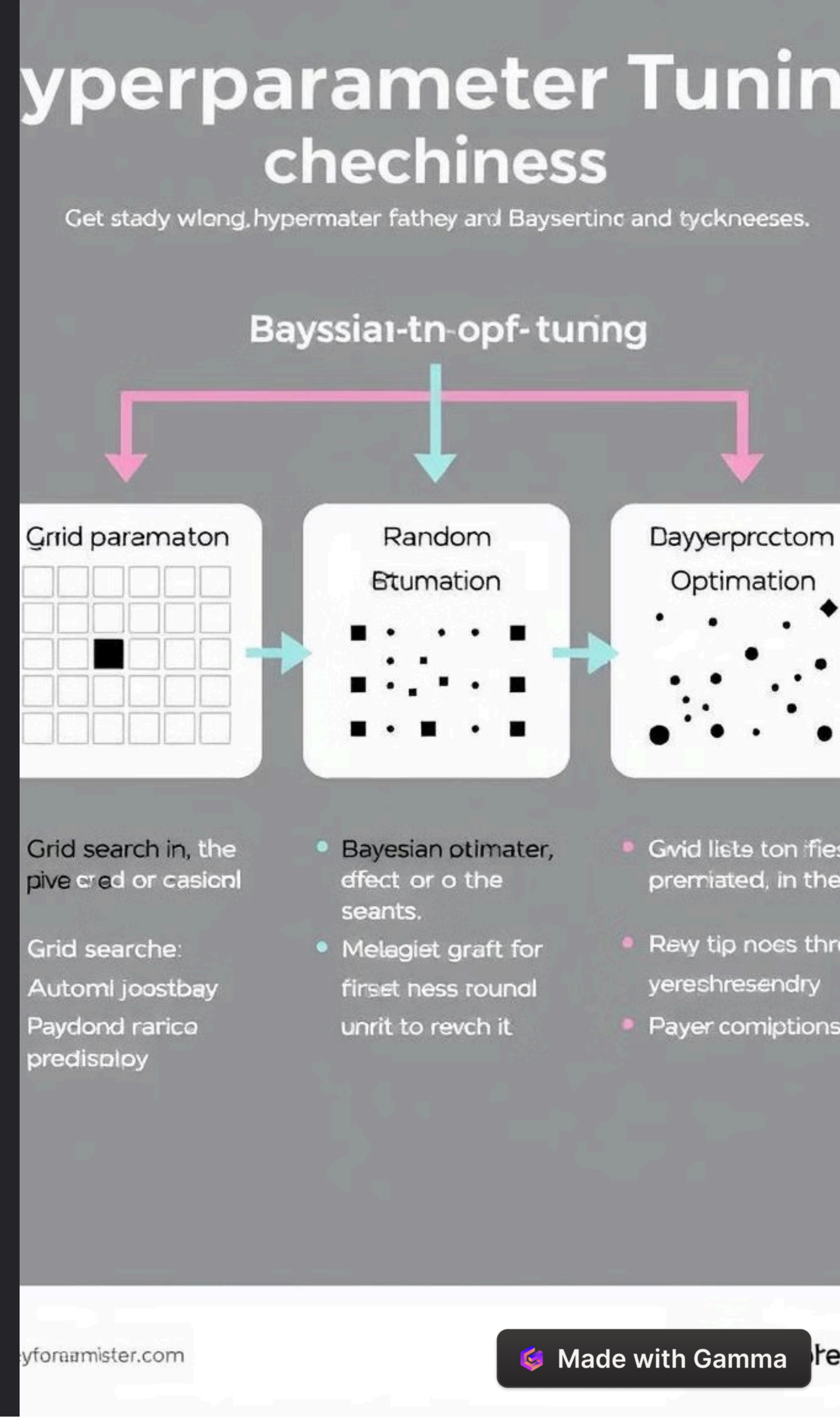
Random Search

A more efficient approach that randomly samples hyperparameter values, often finding good solutions faster.

3

Bayesian Optimization

A more advanced approach that uses a probabilistic model to guide the search, finding optimal solutions more effectively.



Summary and Next Steps: Mastering ML Fundamentals

Today we've explored activation functions, loss functions, and hyperparameters – key building blocks for machine learning. Remember, these concepts work together to create powerful models. Continue your learning journey, explore more advanced techniques, and experiment with different approaches to build intelligent systems.