**Vanishing Gradient 📉**

This happens when gradients become so small they are practically zero. As a result, the model's weights stop updating, and the network

**stops learning** altogether, even if its performance is still poor.

**Exploding Gradient 💥**

This is the opposite issue, where gradients accumulate and become excessively large. This leads to massive, unstable updates to the weights, causing the training process to

**diverge** instead of improving

hyperparameters

are the settings and options for a neural network that you decide on

**before the training process begins**. They control the overall structure of the model and how it learns.

Examples include:

* The

**learning rate**

* The number of

**layers** and **neurons** in the model

* The

**dropout rate**

* The choice of

**optimizer** and **loss function**

Data augmentation

is the process of artificially creating more training data by making modified copies of existing data, like images. This is done by applying transformations such as

**rotating**, **flipping**, or **cropping** the original images.