**Report on Neural Network Optimizer Performance**

This report analyzes the training performance of a custom-built neural network on a synthetic dataset, comparing three different optimization algorithms: Gradient Descent, RMSProp, and Adam. The model was trained for 1,000 epochs, and its performance was measured by loss and accuracy.

**Findings**

The training results clearly demonstrate the superior performance of adaptive optimizers (RMSProp and Adam) compared to the standard Gradient Descent.

* **Gradient Descent:** The model trained with Gradient Descent struggled to converge effectively on this non-linear dataset. It achieved a final accuracy of **0.5000**, indicating it was unable to learn the underlying patterns of the data. This is a common issue with Gradient Descent on complex problems, as it can get stuck in local minima or require very specific, carefully tuned learning rates.
* **RMSProp:** Using an adaptive learning rate, RMSProp demonstrated a significantly better ability to navigate the loss landscape. It achieved a final accuracy of **0.9990**, showing that it successfully learned the data's complex structure.
* **Adam:** The Adam optimizer, which combines the benefits of RMSProp and momentum, performed similarly well, achieving a final accuracy of **1.0000**. The convergence plots confirm that both Adam and RMSProp were able to find an optimal solution quickly.

**Conclusion**

The comparison highlights a critical aspect of training deep learning models: the choice of optimizer has a profound impact on performance. While Gradient Descent is a foundational concept, modern adaptive optimizers like RMSProp and Adam are essential for achieving high accuracy and efficient training on non-linear datasets. They automate the process of finding an effective learning rate for each parameter, making the training process more robust and reliable.

The plots from the code further illustrate this, showing the rapid decrease in loss and the swift increase in accuracy for the Adam and RMSProp models, in stark contrast to the stagnant performance of the Gradient Descent model.