Summary Report

Assignment 2

Advanced Machine Learning_64061

MSBA

Submitted by: Dileep Kumar Pasala

❖ Overview:

To determine how well each neural network model performed on the provided dataset, a total of eleven models were created and assessed. Even though each model used somewhat varied architectures, with variations in regularization, dropout, number of layers, and activation functions, their performance measures were astonishingly constant, suggesting a stable learning environment and a well-tuned dataset.

❖ Model Architecture Summary:

- Models 1 through 3 use ReLU activation and sigmoid output to create two to three dense layers. Efficient, stable, and straightforward.
- Model 4: A single 16-unit hidden layer (ReLU) with a minimalistic design. Performance was comparable to those of deeper models.
- Model 5: Added 32-unit layers and had a little higher capacity, but there was no discernible increase in precision.
- Tanh activation and L2 regularization (λ=0.01) were used in Model 6 to lessen overfitting; the outcomes were comparable to those of ReLU models.
- The most intricate model, Model 7, has several Dense and Dropout layers. No quantifiable improvements, but strong regularization and decent generalization.
- Model 8: A straightforward but efficient Tanh + Dropout combo. performance that is comparable to larger models.
- Model 9: A compact ReLU model with L2 regularization that is reliable and effective, perfect for applications with constrained computing power.
- Dropout layers with moderate hidden sizes were reintroduced in Model 10-11. Model 11 did a good job at balancing performance and simplicity.

❖ Performance Overview:

Metric	Range Observed
Training Accuracy	~98.9%
Validation Accuracy	~87.4%
Test Accuracy	~88.9%
Training Loss	~0.012

Validation Loss	~0.099
Test Loss	~0.086

The findings indicate robust learning behavior and well-optimized hyperparameters since all models performed consistently and well.

❖ Interpretation:

- Accuracy was not improved by adding additional layers or neurons than a moderate depth, suggesting representational sufficiency.
- There was minimal effect of regularization (L2, Dropout), suggesting that there was little overfitting in this dataset.
- Tanh demonstrated no discernible benefit; ReLU activation continuously offered quick convergence and accuracy.
- While more efficient, simpler models (Models 4 and 9) performed on par with more complicated ones.

❖ Key Insights:

- Stable generalization and high accuracy were displayed by all 11 models.
- The simplicity and effectiveness of the Models 4 and 9 make them stand out as the best for deployment.
- When complexity was increased above what was required, the benefits decreased.
- All structures showed consistent training stability.

❖ Conclusion and Recommendations:

The finest overall balance between accuracy and computing efficiency was offered by models 4 and 9. To improve model resilience and convergence, future research may investigate learning rate scheduling, batch normalization, and leaky ReLU or ELU activations.

In summary, all models demonstrated remarkable performance, with deeper designs being as accurate as simpler ones. This supports the fundamental idea that balance and suitable activation functions are frequently more important in neural network design than architectural depth.