Assignment Report - AML Neural Network

Introduction:

This assignment explores the development and optimization of a neural network model for sentiment analysis on the IMDB movie reviews dataset. Both machine learning (ML) and deep learning (DL) techniques were employed to evaluate the model's effectiveness.

Activation Functions Used:

Various activation functions were experimented with, including:

- ReLU (Rectified Linear Unit)
- Sigmoid
- Tanh (Hyperbolic Tangent)

Strategies for Model Enhancement:

To improve the model's performance, several strategies were considered:

⇒ Adding more layers to increase model complexity.
⇒ Modifying activation functions for better non-linear pattern learning
⇒ Fine-tuning the learning rate to enhance convergence.
⇒ Tailoring the loss function to prioritize specific metrics.
⇒ Experimenting with different optimizers for improved convergence
⇒ Leveraging pre-trained models through transfer learning.
⇒ Adjusting batch size to control convergence speed and prevent overfitting.
⇒ Implementing regularization techniques like L1, L2, and dropout to combat overfitting.

Experimental Results:

Hidden Layer Experimentation:

Initially, a model with two hidden layers was employed, resulting in the following outcomes:

- Single-layer approach: Validation loss of 0.2697 and Validation accuracy of 0.8903.
- Three-layered approach: Validation loss of 0.2897 and Validation accuracy of 0.8674.

The number of hidden units was also varied, with the following results:

- Using 32 units: Validation loss of 0.2783 and Validation accuracy of 0.889.
- Using 64 units: Validation loss of 0.2746 and Validation accuracy of 0.8878.

Loss Function and Activation Function Changes:

Adjustments were made to the loss function and activation function:

- Using mean squared error (MSE) loss function: Validation loss of 0.0838 and Validation accuracy of 0.8863.
- Employing tanh activation function in early layers and softmax in the output layer.

Dropout Regularization:

Dropout regularization was introduced to address overfitting:

- Dropout rate of 0.4 in the single-layer model.
- Dropout rate of 0.5 in the three-layer model.

Summary:

For implementing the neural networks, the following layers were used:

- Input Layer: Created using "Keras.Sequential" to represent the input.
- Hidden Layer: Added layers using the format "model.add(Dense(32, activation='tanh'))."
- Output Layer: Typically consisted of a single unit with the sigmoid activation function.

The model was compiled using the Adam optimizer and mean squared error (MSE) as the loss function.

Conclusions

- Neural networks with both single and three layers were explored.
- Tanh activation functions were used instead of ReLU.
- The Adam optimizer was preferred over RMSprop.
- Dropout layers with different rates were introduced.
- L1 and L2 regularizers were incorporated into the models.
- A training accuracy of 99% on the IMDB dataset was achieved.

Comparison of Approaches

Approach	Training Accuracy	Validation Accuracy
Single layer	99.81%	86.34%
Three-layered	99.81%	86.41%
Dropout and regularize	97.44%	88.48%

Observations

- The single-layer approach yielded a validation accuracy of 86.34%.
- The three-layered approach achieved a higher validation accuracy of 86.41%.
- Addressing overfitting with dropout layers resulted in an accuracy increase to 88.48% in the three-layered approach.

Report Conclusion:

This report summarizes the experimentation and findings related to the optimization of neural network models for sentiment analysis on the IMDB dataset. Various strategies were explored, leading to improved model performance and robustness against overfitting. The results demonstrate the impact of different techniques on model accuracy and provide insights into model selection for specific problem domains.