SUMMARY REPORT

1. **Introduction:**

   I used a variety of deep learning models in this assignment to categorize movie reviews as either good or negative using the IMDB sentiment analysis dataset. Investigating the effects of various model architectures, activation functions, and regularization strategies on the model's sentiment prediction was my aim. Finding the model that offers the optimal balance between accuracy and efficiency was the main goal of comparing the performance of simpler and more complicated models.

2. Data Preparation:

To begin, I loaded and preprocessed the IMDB dataset. Every movie review was transformed into a series of numbers, with one number representing each word. The length of each sequence was then adjusted to make sure it was the same. The data could now be sent into the neural network models according to this preprocessing phase.

3. Model Development:

A number of models with various configurations were constructed and tested. The main variables that changed were:

There might be anywhere from one to three hidden layers.

Activation functions: To determine which yielded superior outcomes, I employed Tanh and ReLU.

Dropout layers: I included dropout layers in one of the models to avoid overfitting.

Loss functions: Mean Squared Error and binary cross-entropy were also tested.

Our models included:

- A baseline model with 16 units per layer and two hidden layers.
- A single hidden layer in a single layer model.
- Three hidden levels in a three-layer model.
- A Tanh Activation Model that substitutes the Tanh activation function for ReLU.
- A dropout model with dropout layers to lessen overfitting.

An MSE loss model, in which the loss function is Mean Squared Error rather than binary cross-entropy.

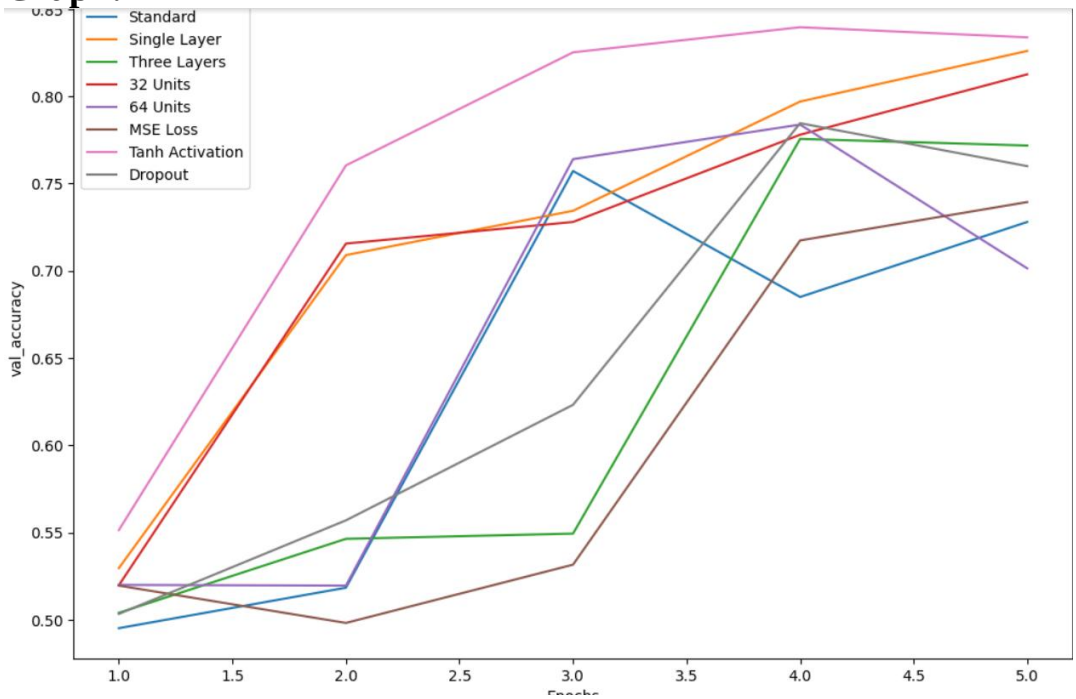
4. **Training**:

A batch size of 512 was used to train each model across five epochs using the training data. In order to assess the model's performance during training and avoid overfitting, I additionally employed a 20% validation split. When the model stopped becoming better, early halting was used to end training.

5. **Evaluation:**

Using a different test set, I assessed each model's performance following training. Accuracy and loss were the two main measures I used to assess the models. These measures gave us insight into each model's accuracy in classifying movie reviews.

6. Graph:



The above graph shows the performance of several neural network models with different configurations, such as MSE Loss, Tanh Activation, Dropout, 32 Units, 64 Units, Single Layer, and Three Layers. It draws attention to the ways that various elements, including regularization strategies, layer counts, and activation functions, affect model performance. The best performance is displayed by the Tanh Activation model(pink line), which has the lowest loss and the highest accuracy. While the Dropout model (gray line) shows variations, indicating instability, the Single Layer model (orange line) likewise works well. The graph demonstrates how layer complexity and activation functions have a major impact on learning outcomes.

7. The models' outcomes on the test set are as follows:

Model	Loss	Accuracy
Standard Model	0.5201	73.10%
Single Layer Model	0.3826	83.07%
Three Layers Model	0.4796	77.45%
Tanh Activation Model	0.3637	83.54%
Dropout Model	0.5397	75.18%

8. In conclusion:

- With the highest accuracy (83.54%) and lowest loss (0.3637), the Tanh Activation Model outperformed ReLU, indicating that tanh activation is more successful in this situation.

- With an accuracy of 83.07%, the Single Layer Model also demonstrated good performance, suggesting that a more straightforward architecture can be very successful.
- Although it did not outperform the single-layer model, the Three Layers Model offered a moderate increase, indicating that performance is not always improved by adding additional layers.
- The Dropout Model's poorer accuracy (75.18%) suggests that there may be too much information loss due to the existing high dropout rate.
- With the lowest accuracy (73.10%), the Standard Model (Baseline) demonstrated that high performance cannot be achieved with a simple design without optimization.

9. **Suggestions:**

To get the greatest results, use the Tanh Activation Model.
If computational efficiency is important, take a look at the Single Layer Model.
Adjust the dropout rate to strike a balance between information retention and regularization.
It's not always helpful to add more layers; instead, concentrate on activation functions and layer optimization.
Model performance may be enhanced by additional tuning (e.g., batch normalization, other optimizers).
In summary, the choice of activation function is critical to model performance, and when appropriately tuned, simpler models can frequently outperform deeper designs.