**Summary of the model:**

1. The graph indicates that Model Hyper was used to attain the highest accuracy and loss. We used three 32, 32, and 16-thick layers, each with a 0.5 dropout rate. To compare the model's MSE loss to any other model, we use MSE (other than Final Hypertuned Model) We see that the loss is minimal while using MSE.
2. Units: We can see that the accuracy varies very little while the loss function reduces when we increase the unit from 16, 32, 64, and 128.
3. We discovered that increasing the number of hidden layers from 1 to 3 had no impact on the precision or deflection of the loss function.
4. Loss Function: "mse" produces the lowest loss value when employed as the loss function for the IMDB dataset in comparison to binary cross entropy.
5. Activation Function: Because of the vanishing gradient issue, the model's tanh activation function performs poorly.
6. Regularization: In comparison to the initial model, regularization produces a significantly lesser loss with less overfitting, with the L-2 model displaying a marginally higher accuracy.
7. Dropout: The dropout does not affect accuracy; however it does help to reduce the loss function.