**Dataset Selection:**

I chose the Iris dataset for this task. The Iris dataset is a simple and widely used dataset for classification. It consists of features related to iris flowers and corresponding labels indicating the iris species.

**Basic Neural Network:**

I created a simple neural network using TensorFlow. The architecture includes an input layer, a hidden layer with a ReLU activation function, and an output layer with softmax activation for multi-class classification. The neural network is designed as follows:

- Input Layer: Features

- Hidden Layer: Dense layer with ReLU activation

- Output Layer: Dense layer with softmax activation

**Gradient Descent with Momentum:**

I implemented Gradient Descent with Momentum during the training process. The momentum parameter was set to 0.9. I observed and compared the model's convergence speed and loss reduction with and without momentum. The model trained with momentum achieved faster convergence and lower loss compared to the model without momentum.

**Learning Rate Decay:**

I applied a learning rate decay strategy during training. I started with a higher learning rate and reduced it over epochs. The learning rate decay was implemented using the ExponentialDecay schedule with a decay rate of 0.9. This strategy helps control the learning rate during training and can lead to better convergence. The model trained with learning rate decay showed effective convergence with a slightly slower decrease in learning rate.

**Dropout Implementation:**

I integrated a dropout layer in the hidden layer of the neural network. Dropout is a regularization technique that randomly drops a fraction of neurons during training. I trained the model both with and without dropout and compared their performances on a validation dataset. The model with dropout showed improved generalization on the validation dataset compared to the model without dropout, indicating the effectiveness of dropout in preventing overfitting.

**Training and Evaluation:**

I trained the neural network using the chosen techniques: Gradient Descent with Momentum, Learning Rate Decay, and Dropout. The training process involved fitting the model to the training data for multiple epochs. After training, I evaluated the models on a separate test dataset to measure accuracy and loss.

**Documentation:**

The code is documented with comments to explain each part of the implementation. Observations and results are presented in a clear format, highlighting the impact of each technique on the model's performance. Performance metrics such as accuracy and convergence speed are compared for models with and without each technique.

**Summary:**

- Gradient Descent with Momentum accelerated convergence and reduced loss.

- Learning Rate Decay controlled the learning rate during training, contributing to effective convergence.

- Dropout improved generalization on the validation dataset, preventing overfitting.

The documentation and observations provide insights into the effectiveness of each technique in enhancing the neural network's performance on the Iris dataset.

