Training

We see that setting the learning rate to 0.1 leads to very bad results. This is because the model trains too quick and is unable to get better. We know from this that the learning rate is too high.

When learning rate is 0.01, we get much better results but there are still some fluctuations which doesn't let the model get better in the later stages of training. So, we know that the learning rate is still a little too high.

When the learning rate is 0.001, we get the highest accuracy and see that the model continues to get better even at the later stages. So, this is the best learning rate.

For learning rate equal to 0.0001, the model is slower to train and is still improving when training is stopped. So, this learning rate is a too low.