10000/10000 [==============================] - 3s 319us/sample - loss: 0.0198 - acc: 0.9007

The testing accuracy metric for 0.1 momentum is [0.019832130385935315, 0.9007]

---------------------------------------------------

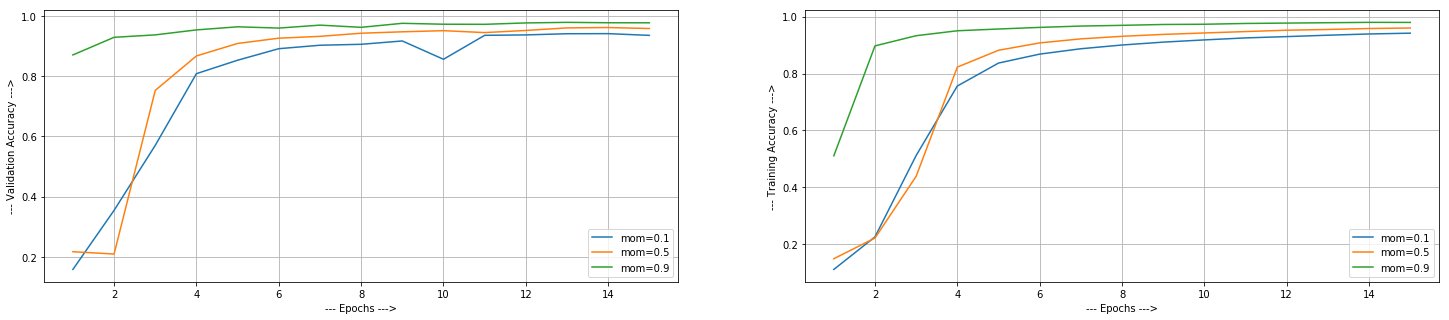
10000/10000 [==============================] - 3s 343us/sample - loss: 0.0096 - acc: 0.9521

The testing accuracy metric for 0.5 momentum is [0.009554959009696317, 0.9521]

---------------------------------------------------

10000/10000 [==============================] - 3s 330us/sample - loss: 0.0048 - acc: 0.9758

The testing accuracy metric for 0.9 momentum is [0.004812074553671141, 0.9758]



**OBSERVATION:**

We saw that as weincreased the momentum parameter, all three validation, test and training accuracy of the model increased maintaining equal number of epochs. This is because SGD with momentum is method which helps accelerate gradients vectors in the right directions, thus leading to faster converging. By observing above graphs, we can say that the model with higher momentum is converging very quickly.