**Session 5**

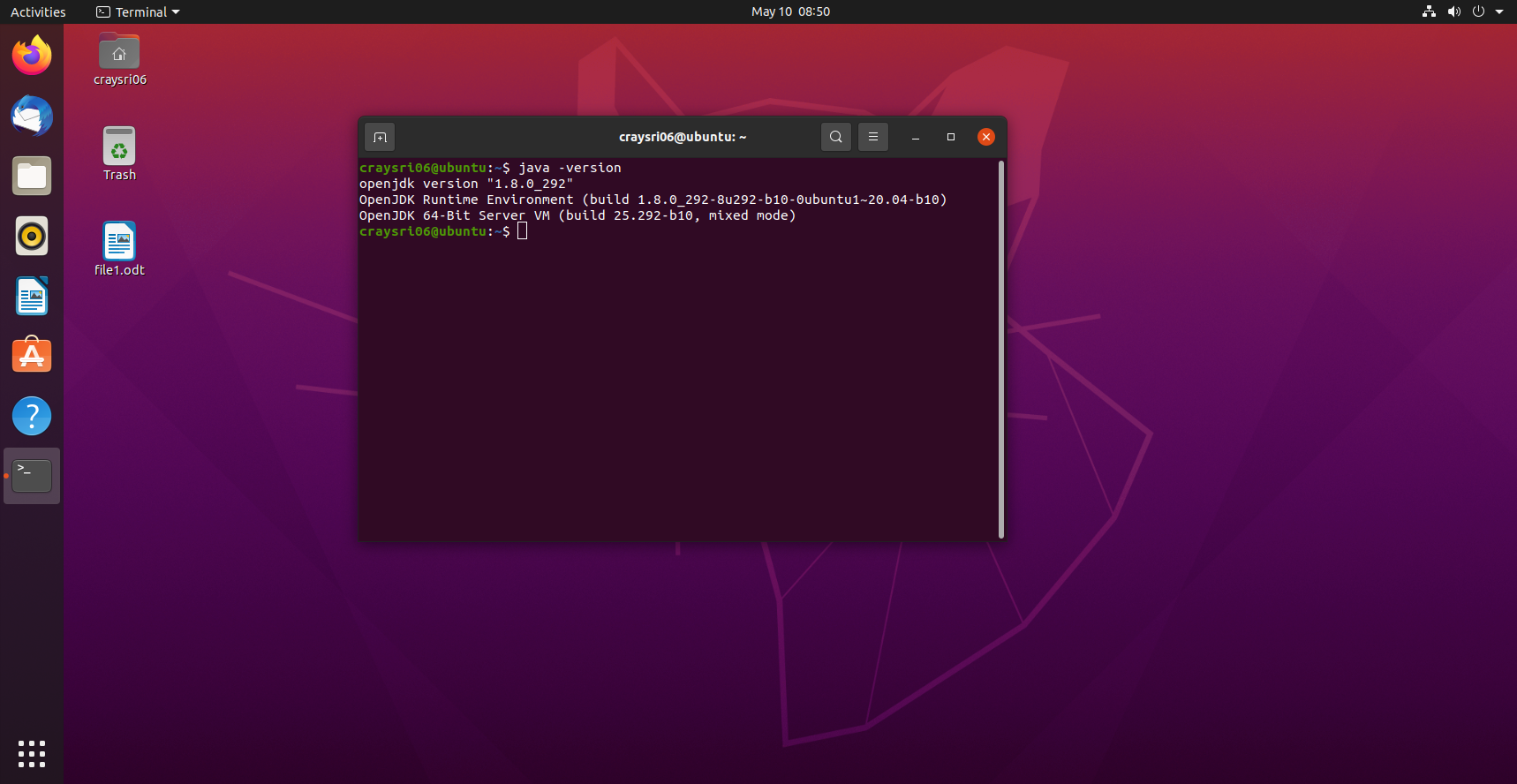
**10/05/2021**

**Problem Statement:**

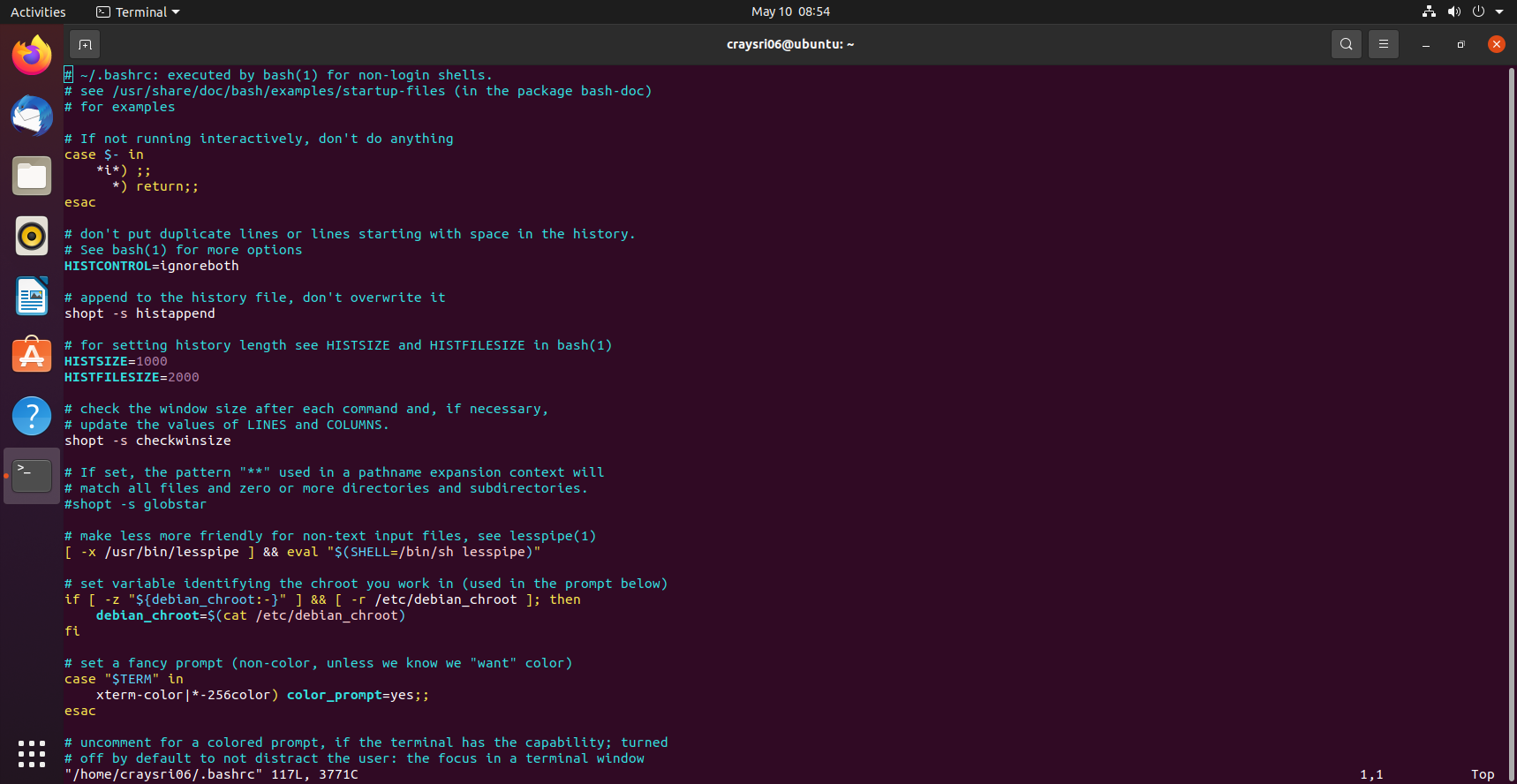
**BD- Session Record- Single Node Hadoop Cluster Setup files**

Different modes of HADOOP

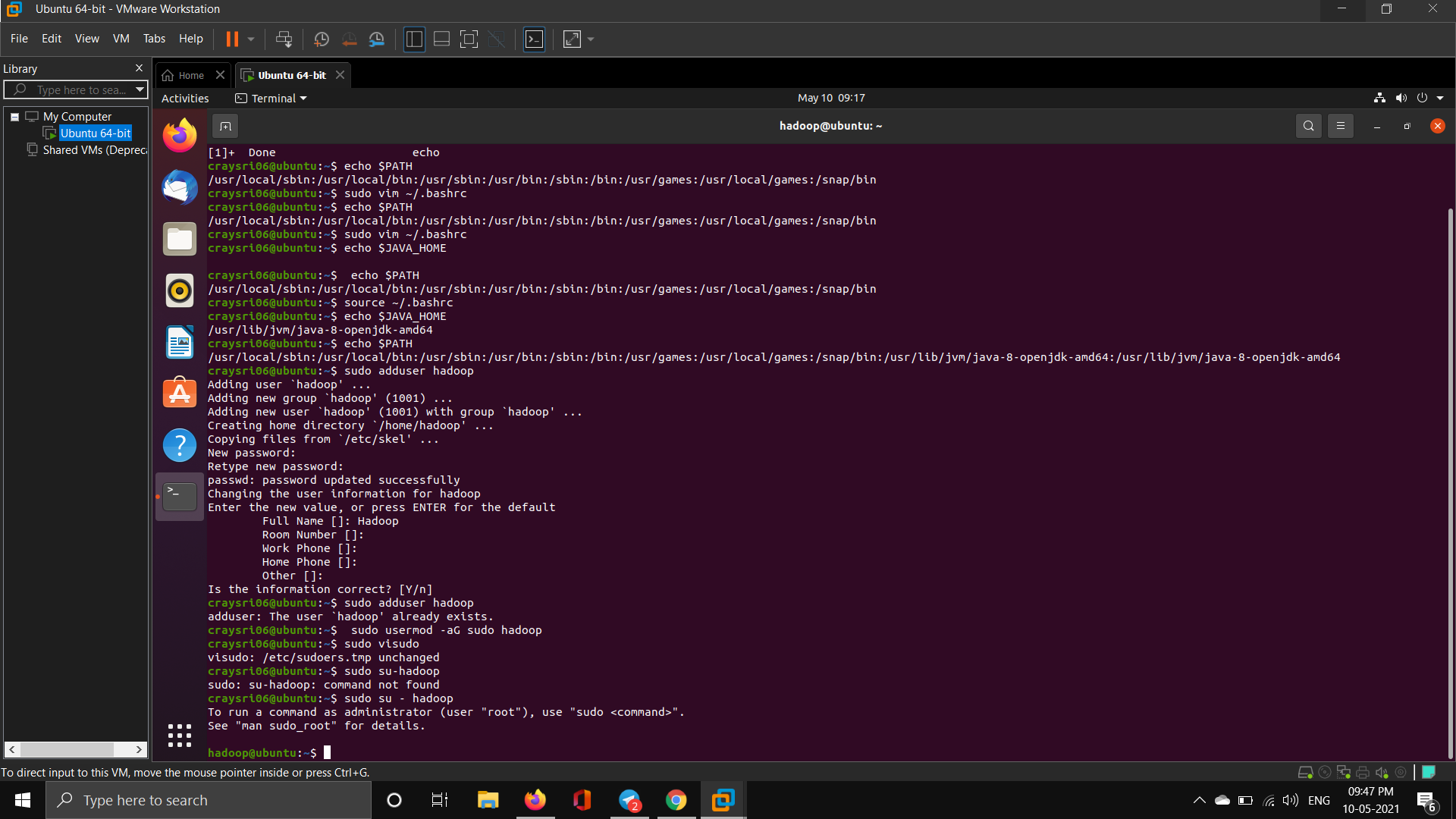
1. Standalone:- It is local to the PC,Cannot connect to the distributed network(cluster),cannot have the connection to the other nodes in network
2. Pseudo-distributed:- Single node cluster,It is reusable & can be used by the other nodes
3. Full-distributed:-It is used for different clusters

**Installation of Java 8 and vim**

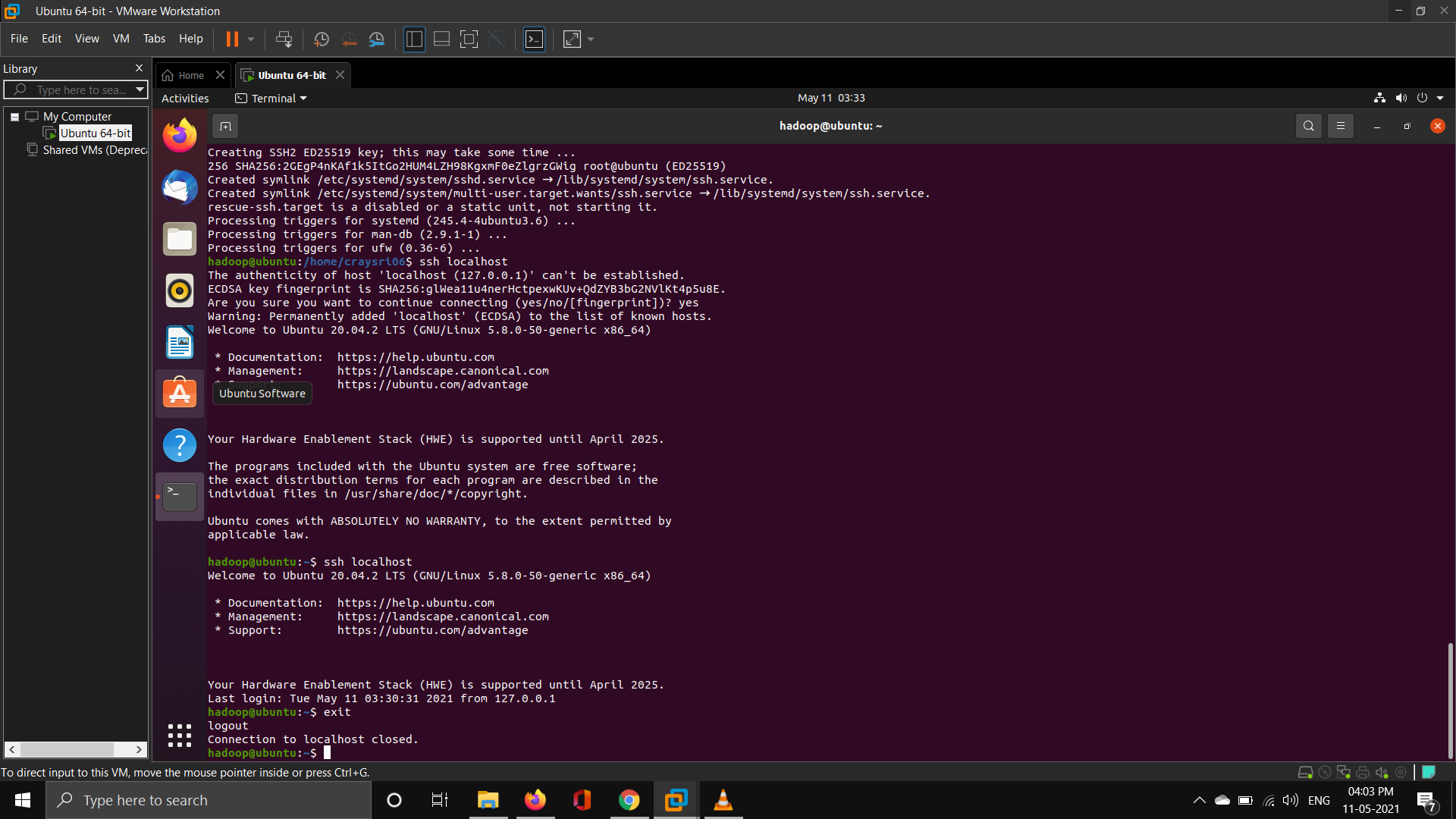
**Adding the Jdk path to the path variable and Inform the OS about the modification**



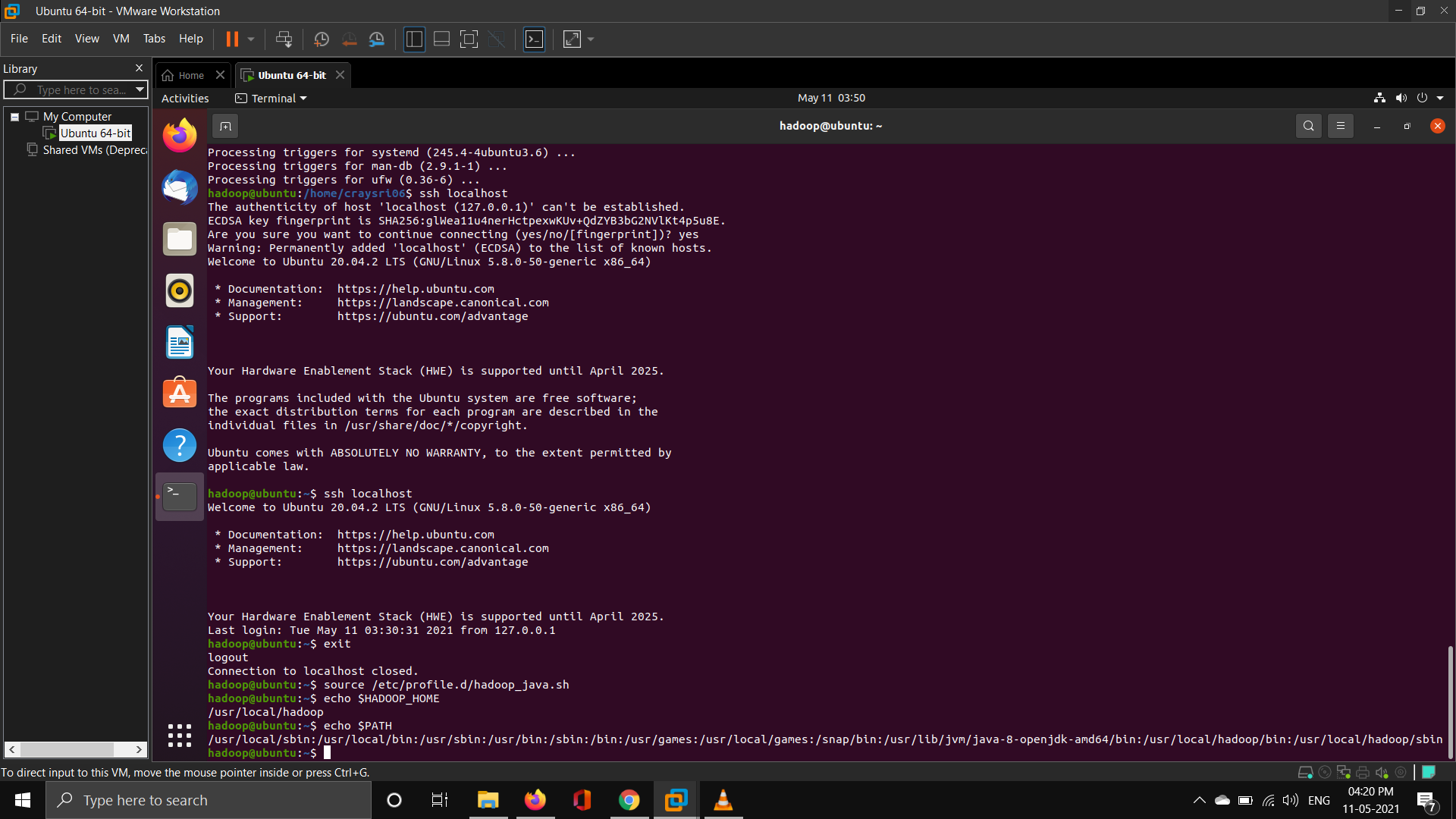
**Add a dedicated user for the HADOOP and switch to that user**



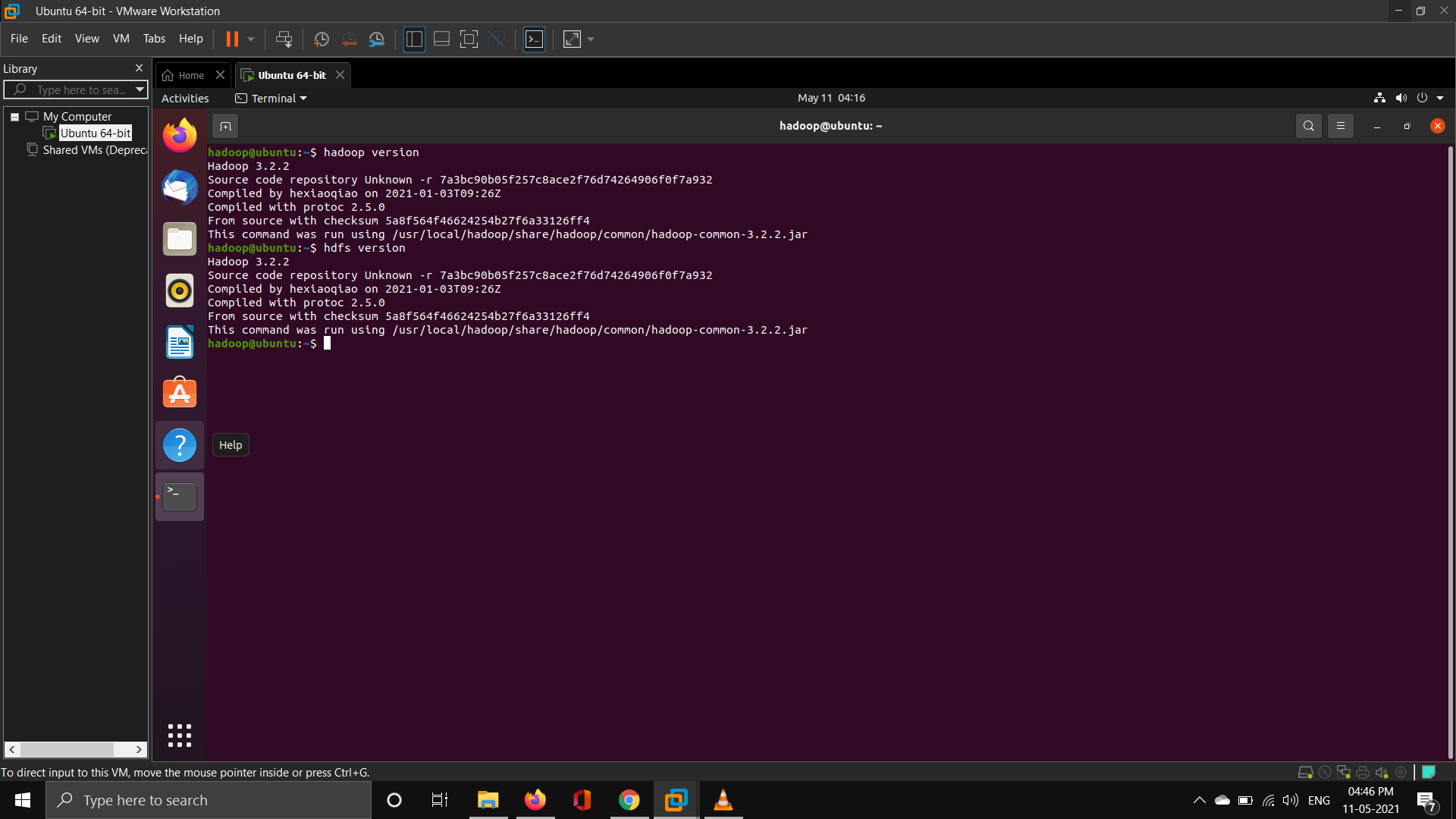
**Once the user is added, login to the user “Hadoop” to generate the ssh key for passwordless login ( hadoop@machinename)**



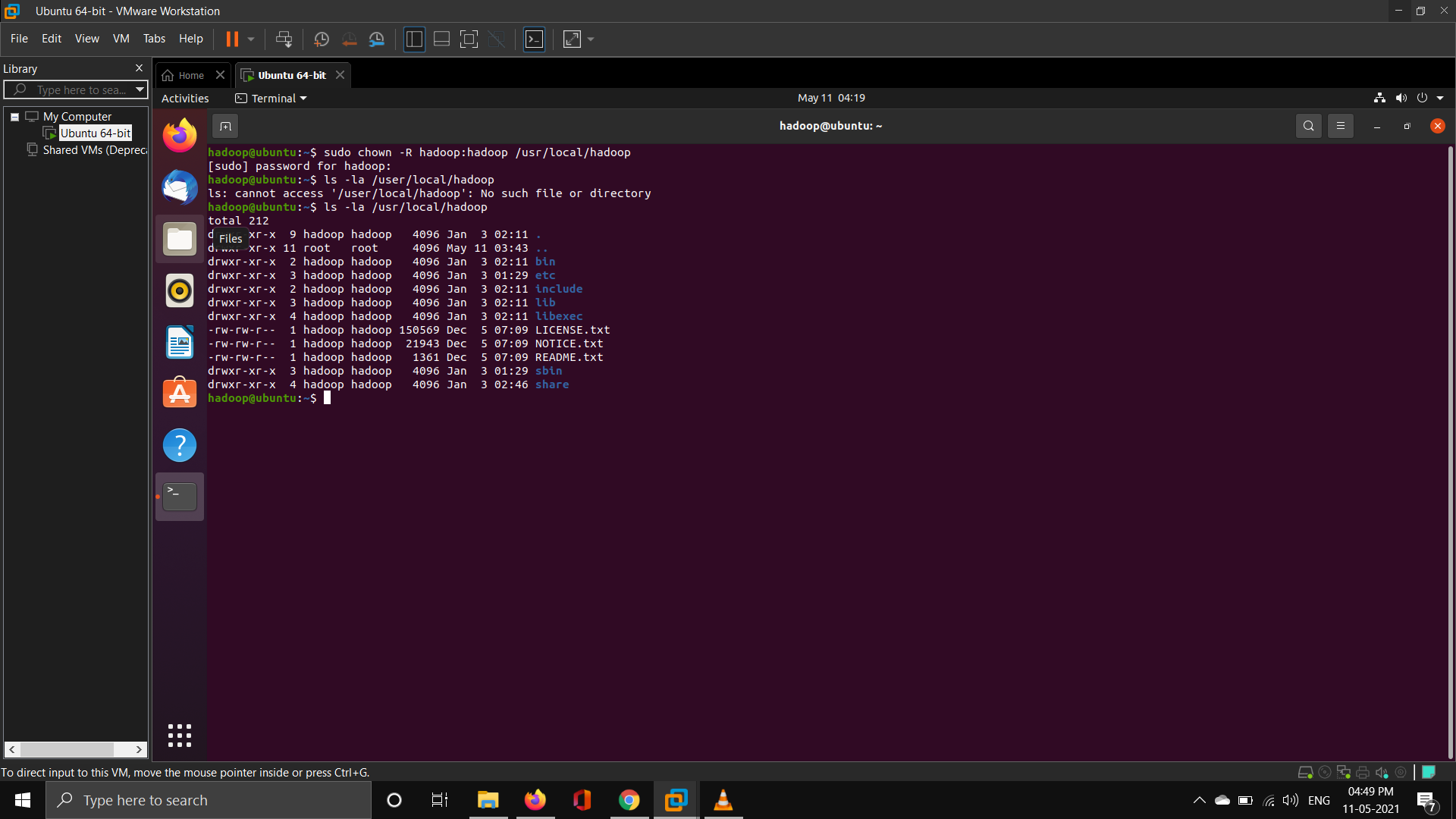
**Setup the path variables for hadoop**



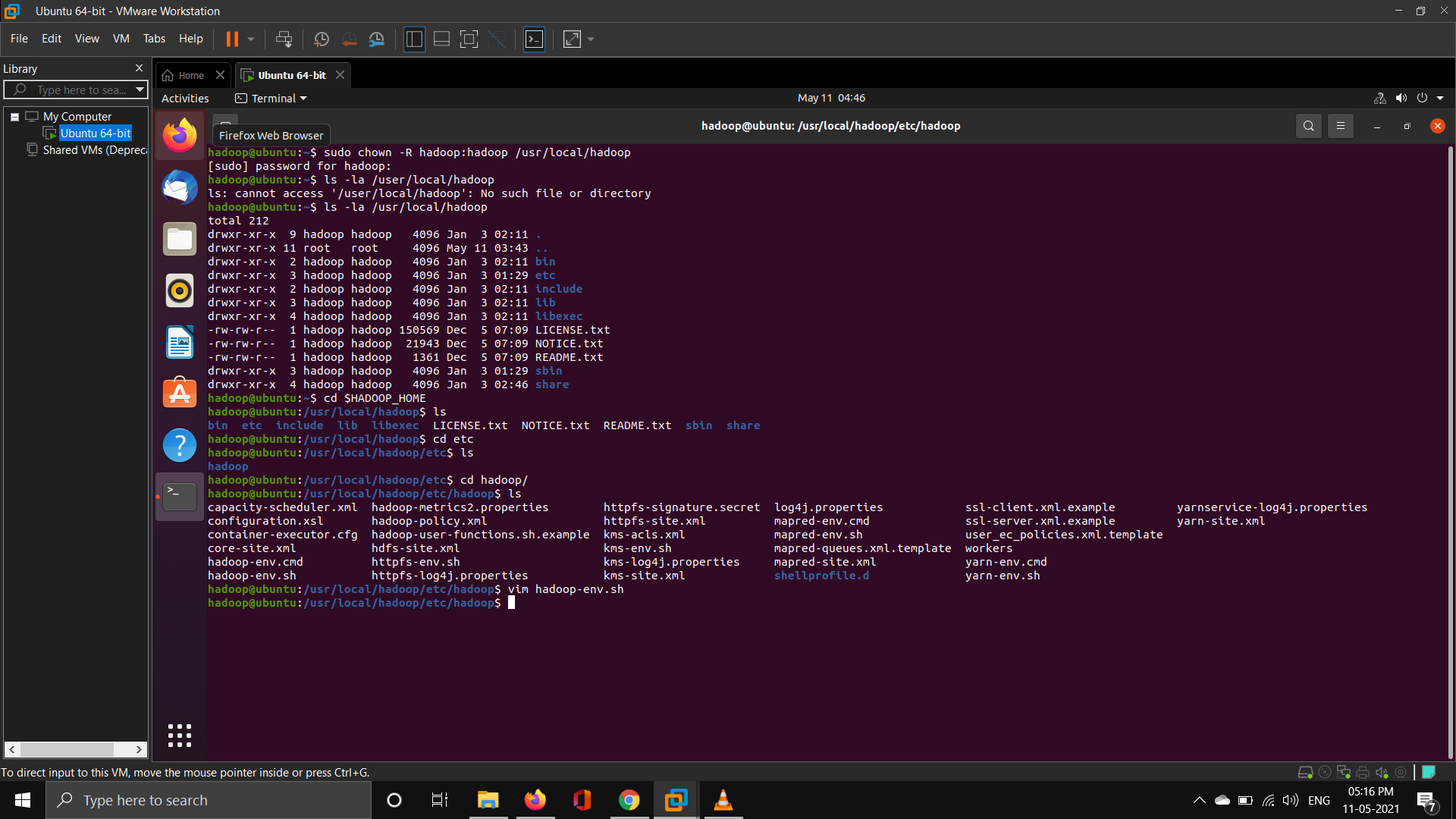
**Confirm your hadoop and hdfs version**



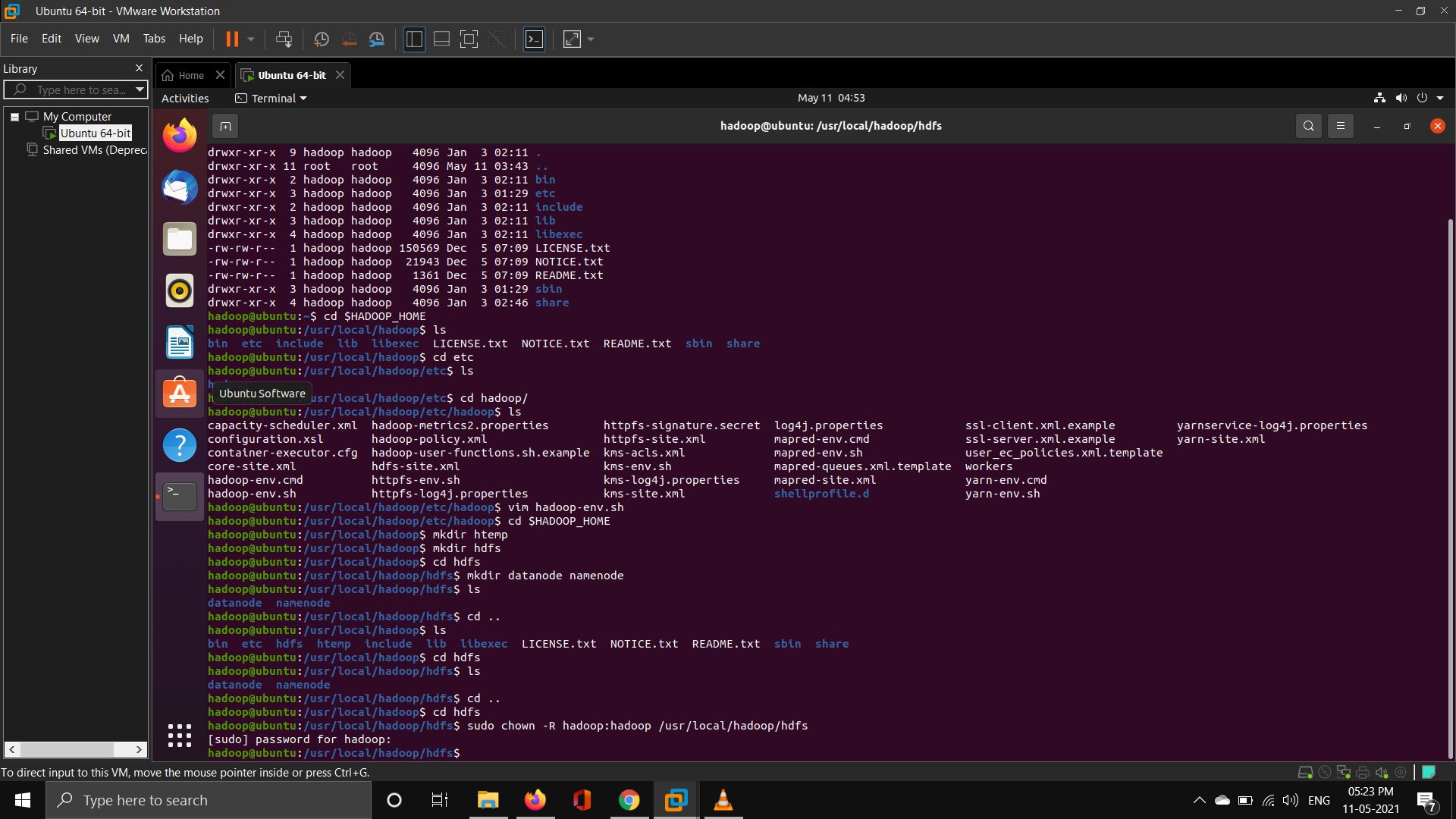
**Give the permission for the hadoop folder to hadoop user**



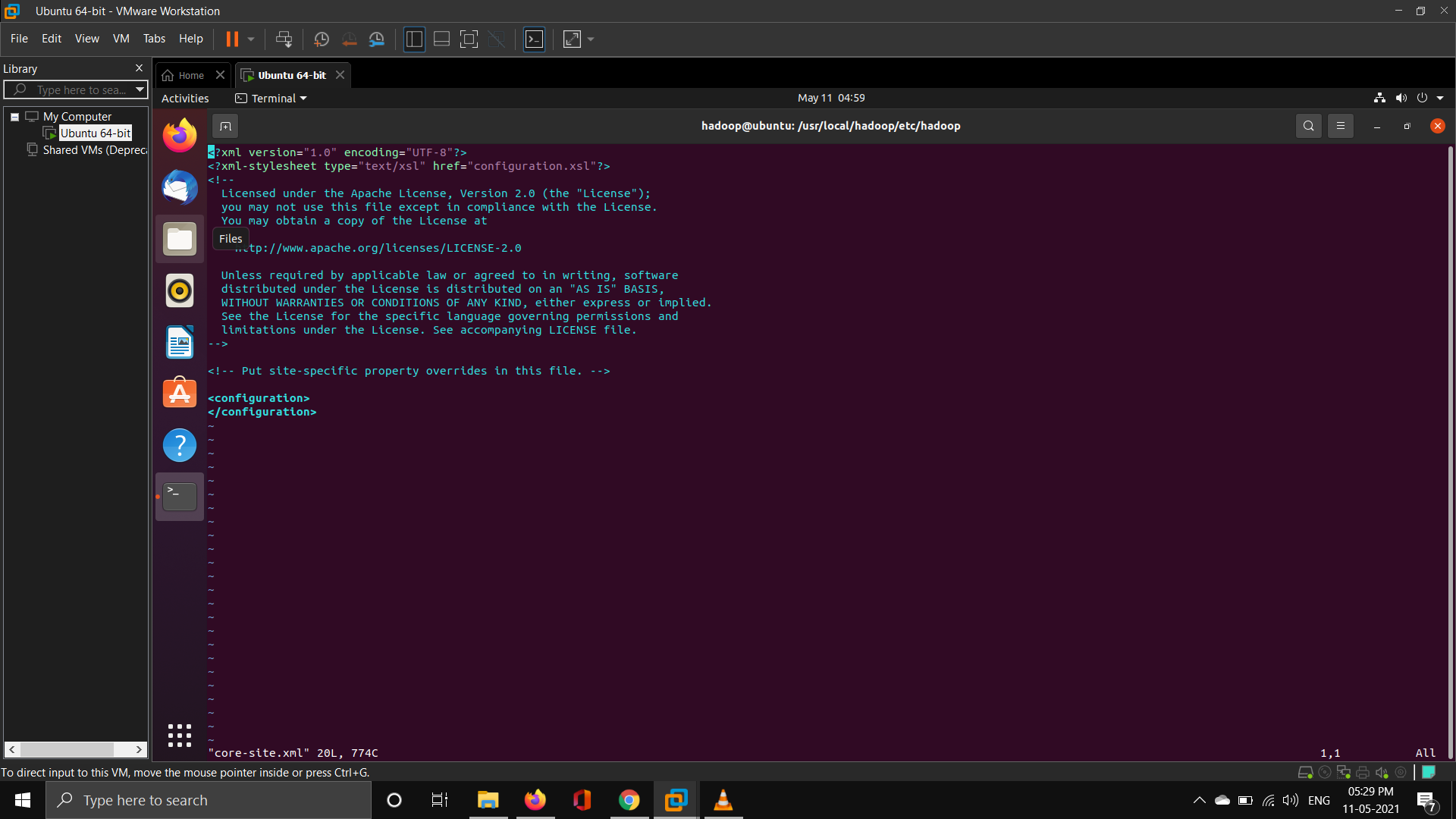
**Specify JAVA\_HOME in hadoop-env.sh (/usr/local/hadoop/etc/hadoop)**



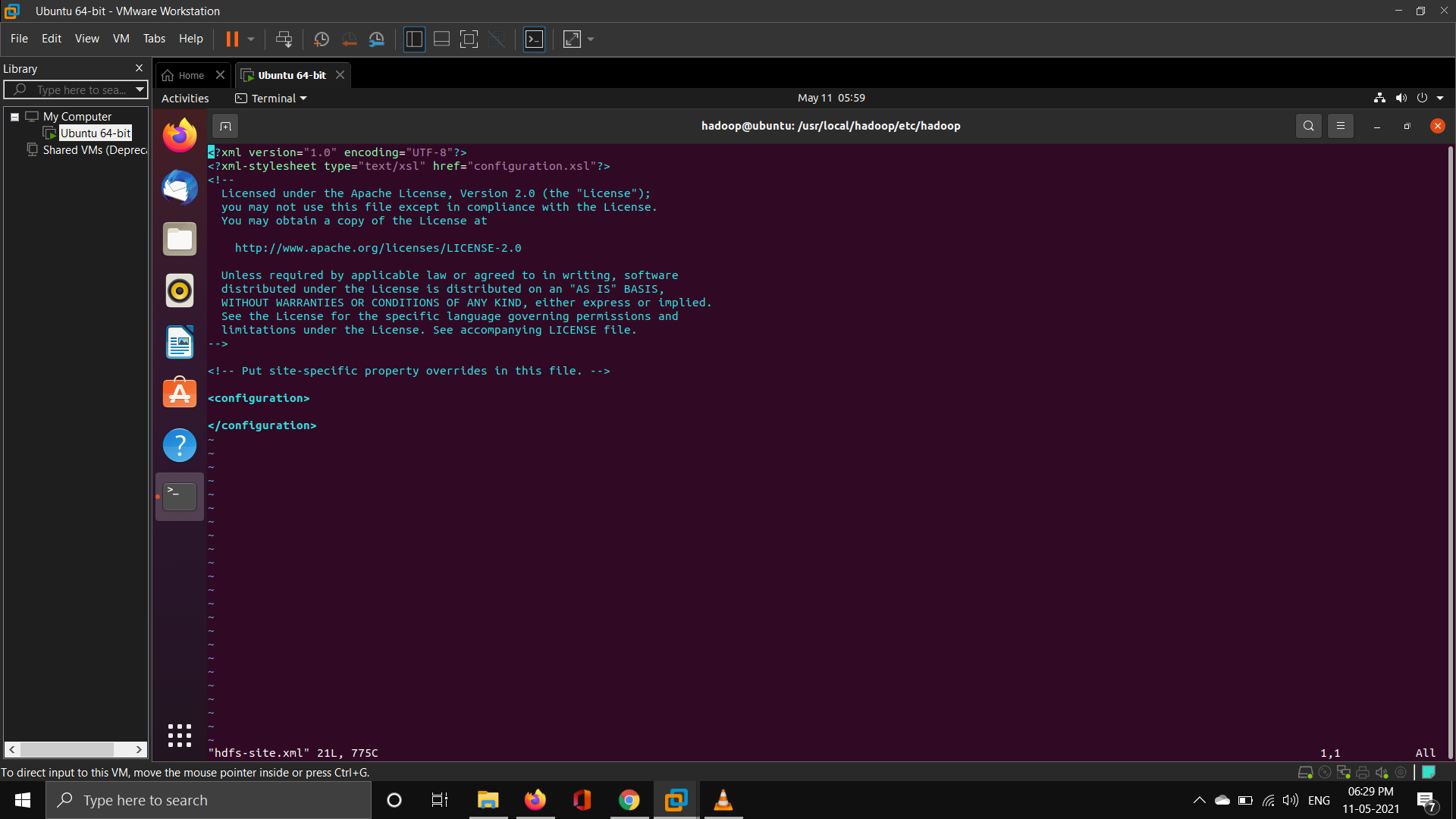
**Give the permission for the hdfs and htemp folder to hadoop user**



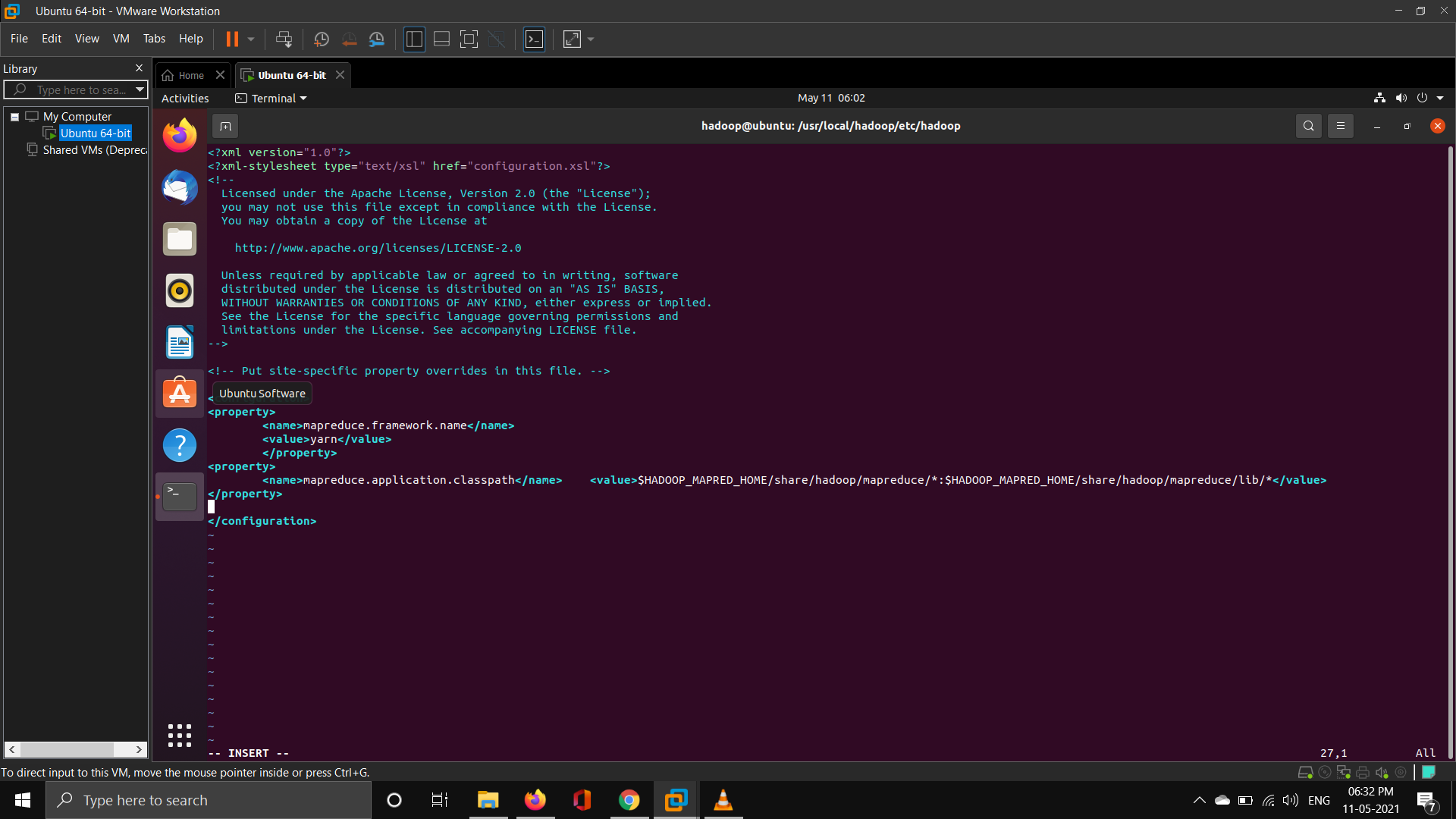
**Modify core-site.xml to setup web portal for hadoop**



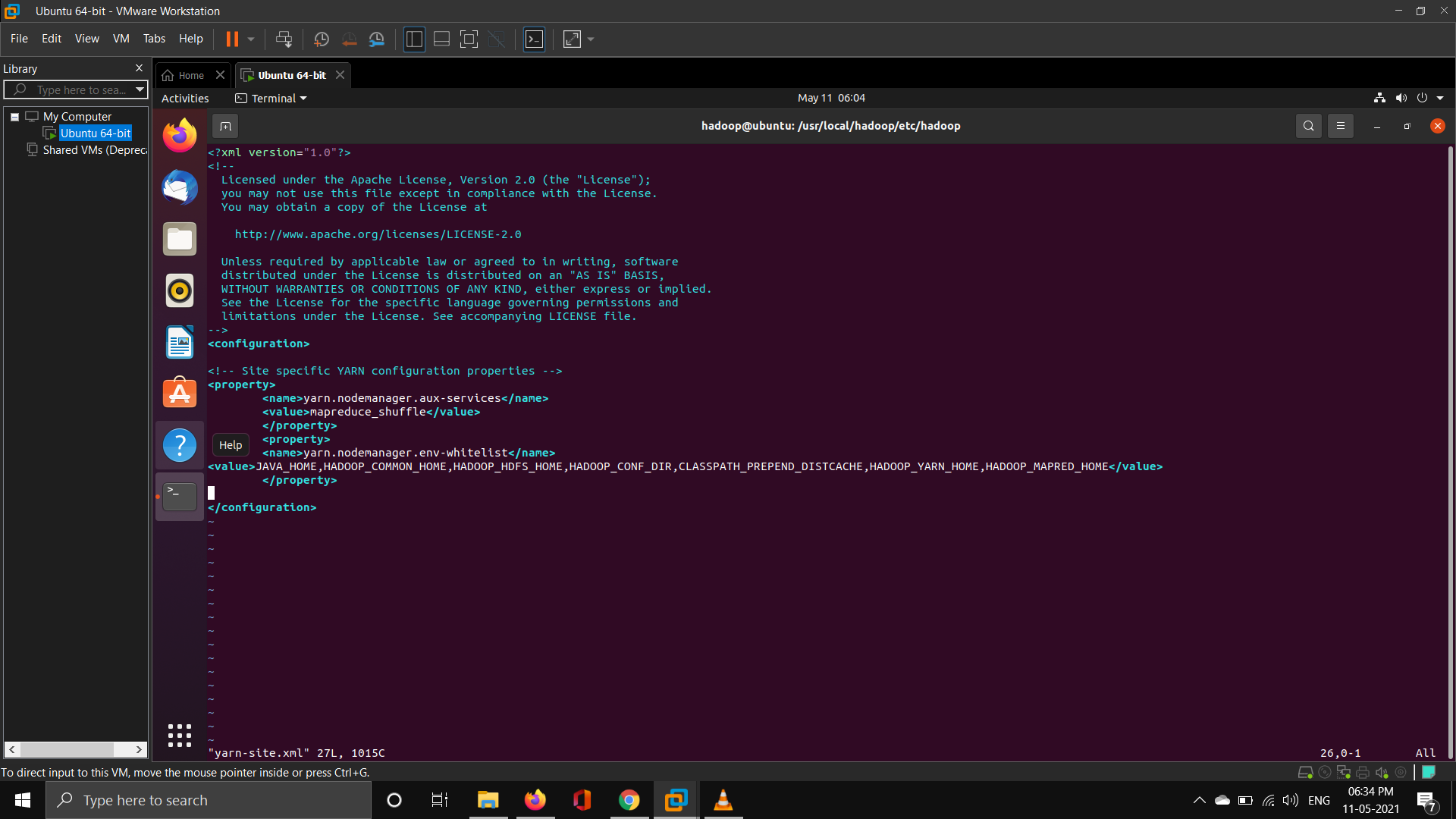
**Modify hdfs-site.xml to setup namenode and datanode path and replication factor**



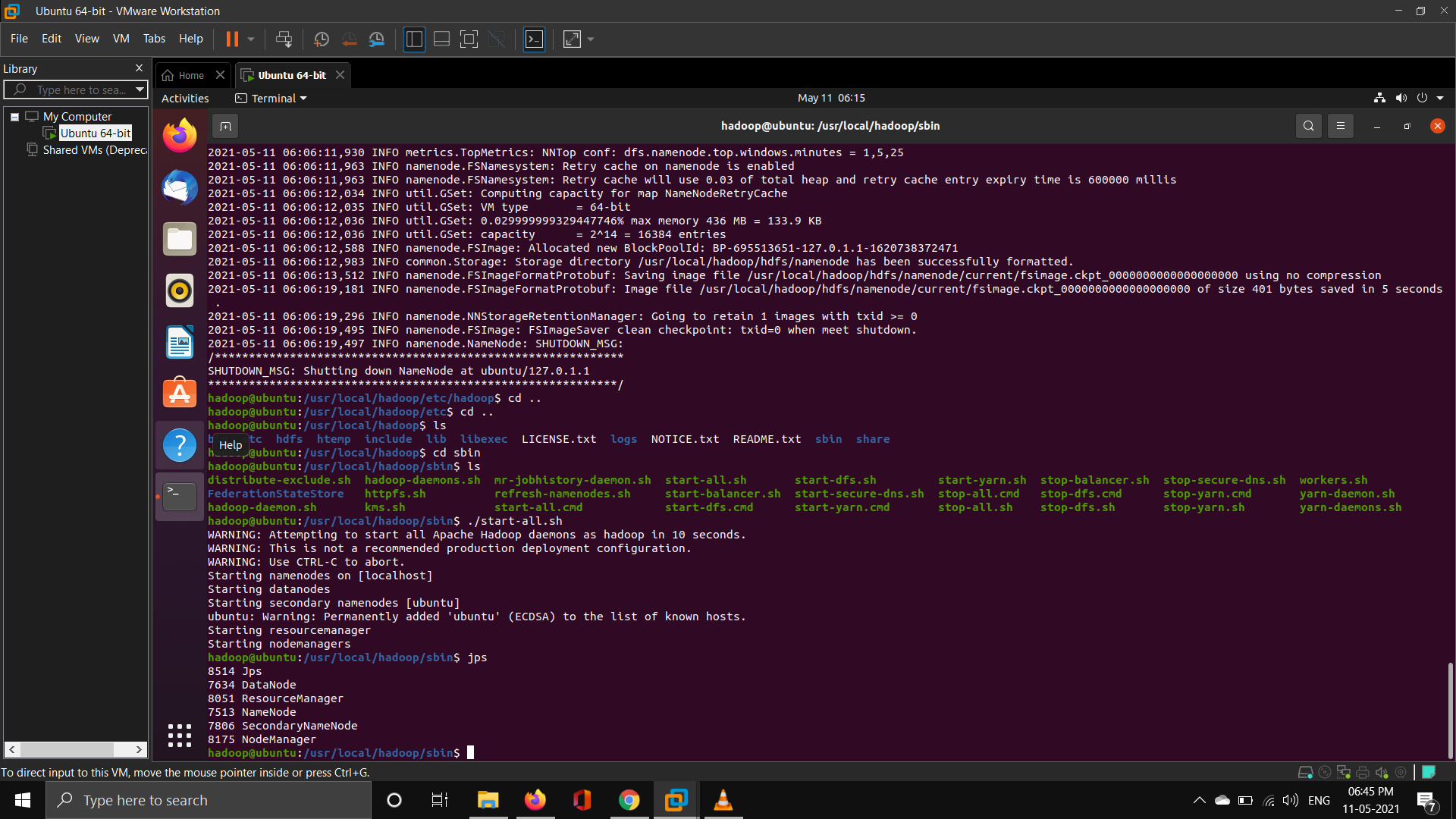
**Configure the mapreduce framework by editing the mapred-site.xml**



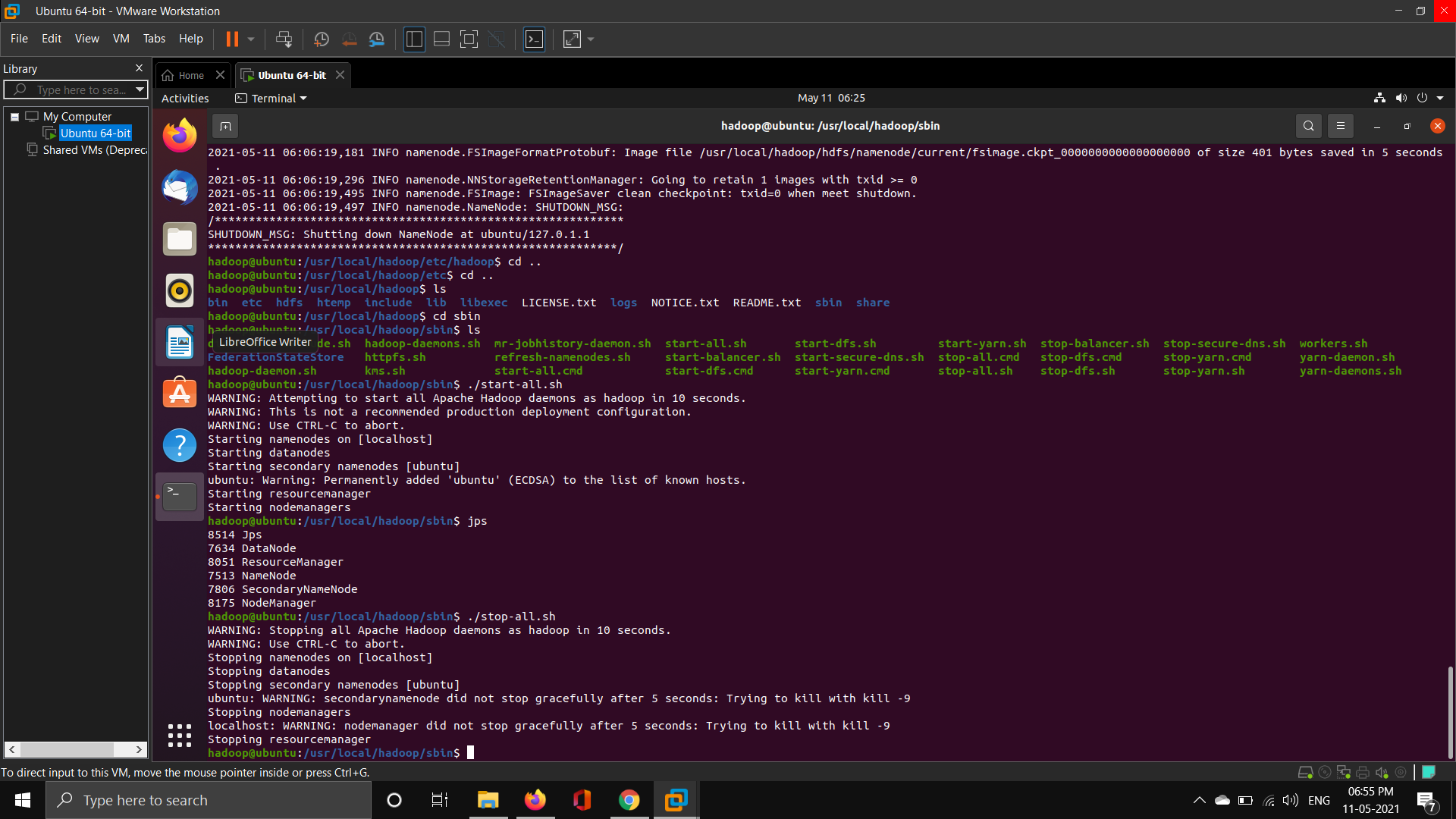
**Configure the YARN resource manager by editing the yarn-site.xml**



**Starting hadoop daemons**



**Stop all daemons**



**Sriranjan S**

**1NT18IS159**