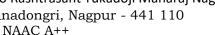




Yeshwantrao Chavan College of Engineering

(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University) Hingna Road, Wanadongri, Nagpur - 441 110





Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration. Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problemsolving skills through emerging technologies.

Session 2025-2026

Vision: Dream of where you want.	Mission: Means to achieve Vision

Program Educational Objectives of the program (PEO): (broad statements that describe the professional and career accomplishments)

PEO1	Preparation	P: Preparation	Pep-CL abbreviation
PEO2	Core Competence	E: Environment	pronounce as Pep-si-lL
		(Learning Environment)	easy to recall
PEO3	Breadth	P: Professionalism	
PEO4	Professionalism	C: Core Competence	
PEO5	Learning	L: Breadth (Learning in	
	Environment	diverse areas)	

Program Outcomes (PO): (statements that describe what a student should be able to do and know by the end of a program)

Keywords of POs:

Engineering knowledge, Problem analysis, Design/development of solutions, Conduct Investigations of Complex Problems, Engineering Tool Usage, The Engineer and The World, Ethics, Individual and Collaborative Team work, Communication, Project Management and Finance, Life-Long Learning

PSO Keywords: Cutting edge technologies, Research

"I am an engineer, and I know how to apply engineering knowledge to investigate, analyse and design solutions to complex problems using tools for entire world following all ethics in a collaborative way with proper management skills throughout my life." to contribute to the development of cutting-edge technologies and Research.

Integrity: I will adhere to the Laboratory Code of Conduct and ethics in its entirety.

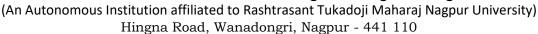
Name and Signature of Student and Date

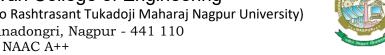
(Signature and Date in Handwritten)





Yeshwantrao Chavan College of Engineering







Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration. Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problemsolving skills through emerging technologies.

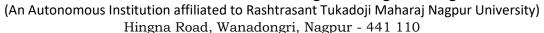
Session	2025-26 (ODD)	Course Name	BIG DATA AND HADOOP-LAB
Semester	7 AIDS	Course Code	22ADS704
Roll No	03	Name of Student	Debasrita Chattopadhyay

Practical Number	01	
Course Outcome	 Understand big data analytics and its business applications. Analyze the HADOOP and Map Reduce technologies associated with big data analytics. Apply Big Data analytics Using Pig and Hive. 	
Aim	Installation of Apache Hadoop on Linux System.	
Problem Definition	Installation of Apache Hadoop on Linux System.	
Theory (100 words)	Apache Hadoop is a free software framework for distributed storage and processing of large sets of data using the MapReduce programming model. It works on a cluster of commodity hardware with a high degree of scalability and fault-tolerance. The Hadoop ecosystem primarily consists of HDFS (Hadoop Distributed File System) for distributed storage and YARN (Yet Another Resource Negotiator) for resource management and job scheduling. To install Hadoop on a Linux system you have to install Java and enable SSH and also modify the core configuration files for pseudo-	
Procedure and	distributed mode or fully distributed mode.	
Execution (100 Words)	 Steps of Implementation Update the Linux system packages using sudo apt update. Install Java Development Kit (JDK) using sudo apt install openjdk-11-jdk. Verify Java installation with java -version. Create a Hadoop user using sudo adduser hadoop. Switch to the Hadoop user with su - hadoop. Configure passwordless SSH using ssh-keygen -t rsa -P "" and cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys. Download Hadoop from the official Apache website using wget. Extract the Hadoop tar file using tar -xzf hadoop-x.y.z.tar.gz. Move Hadoop to /usr/local/hadoop using sudo mv hadoop-x.y.z /usr/local/hadoop. Set environment variables in .bashrc for JAVA HOME 	

Nagar Yuwak Shikshan Sanstha's



Yeshwantrao Chavan College of Engineering





NAAC A++

Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: www.ycce.edu

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

- 11. Edit configuration files (core-site.xml, hdfs-site.xml, mapred-site.xml, yarn-site.xml) in the Hadoop etc/hadoop directory.
- 12. Format the Hadoop namenode using hdfs namenode format.
- 13. Start HDFS and YARN services using start-dfs.sh and start-yarn.sh.
- 14. Verify installation using jps command to check running daemons.
- 15. Access Hadoop web interface at http://localhost:9870/.

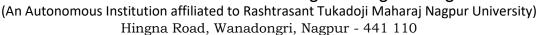
Code:

- # 1. Update system packages sudo apt update -y
- # 2. Install Java sudo apt install openjdk-11-jdk -y
- # 3. Verify Java installation java -version
- # 4. Create a Hadoop user sudo adduser hadoop
- # 5. Switch to Hadoop user su hadoop
- # 6. Configure passwordless SSH ssh-keygen -t rsa -P "" cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys chmod 600 ~/.ssh/authorized_keys ssh localhost # test SSH
- # 7. Download Hadoop (replace version if needed) wget https://downloads.apache.org/hadoop/common/hadoop-3.3.6/hadoop-3.3.6.tar.gz
- # 8. Extract Hadoop tar -xzf hadoop-3.3.6.tar.gz
- # 9. Move Hadoop to /usr/local directory sudo mv hadoop-3.3.6 /usr/local/hadoop

Nagar Yuwak Shikshan Sanstha's



Yeshwantrao Chavan College of Engineering





NAAC A++ Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: <u>www.ycce.edu</u>

Department of Artificial Intelligence & Data Science

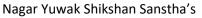
Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

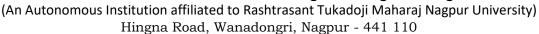
To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

	# 10. Set environment variables nano ~/.bashrc	
	Output: theia@theiadocker-srita201326:/home/project\$ hadoop version bash: hadoop: command not found theia@theiadocker-srita201326:/home/project\$ git clone https://github.com/big-data-europe/docker-hadoop.git cloning into 'docker-hadoop' remote: Enumerating objects: 539, done. remote: Counting objects: 100% (189/189), done. remote: Total 539 (delta 169), reused 166 (delta 166), pack-reused 350 (from 1) Receiving objects: 100% (39/39), 108.00 kiB 13.50 MiB/s, done. Resolving deltas: 100% (35/359), 108.00 kiB 13.50 MiB/s, done. Resolving deltas: 100% (35/359), 108.00 kiB 13.50 MiB/s, done. AMARN(00000) /home/project/docker-hadoop theia@theiadocker-srita201326:/home/projects/docker-hadoop theia@theiadocker-srita201326:/home/project/docker-hadoop d, please remove it to avoid potential confusion [+] Running 28/28 / historyserver Pulled / datanode Pulled / resourcemanager Pulled / resourcemanager Pulled / resourcemanager Pulled / namenode Pulled	3s 3s 3s 3s 7s
	Volume "docker-hadoop hadoop namenode" Created √ Volume "docker-hadoop hadoop_datanode" Created √ Volume "docker-hadoop hadoop historyserver" Created √ Volume "docker-hadoop hadoop historyserver" Created √ Container resourcemanager √ Container datanode thia@theiadocker-srita201326:/home/project/docker-hadoop\$ docker ps	9s 9s
	CONTAINER ID IMAGE STATUS NAMES 1ed7bc3d075c bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8 Up 10 seconds (health: starting) 9864/tcp 7/entrypoint.sh /run." 25 seconds ago	
	datanode 74e6b9be4c20 bde20e/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run_" 25 seconds ago Up 11 seconds (health: starting) 8042/tcp nodemanager 57b2ffa5d876 bde20e/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run_" 25 seconds ago	Ш
	Up 11 seconds (health: starting) resourcemanager fe222273dbse bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run" 25 seconds ago Up 11 seconds (health: starting) 0.0.0.9000->9000/tcp, [::]:9000->9000/tcp, 0.0.0.9870->9870/tcp, [::]:98 0->9870/tcp namenode 13388eebda32 bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run" 25 seconds ago Up 11 seconds (health: starting) 8188/tcp historyserver theia@theiadocker-arita201326:/home/project/docker-hadoop\$ docker exec -it namenode bash root@fe22273db8e:/#	87
	historyserver theia@theiadocker-srita201326:/home/project/docker-hadoop\$ docker exec -it namenode bash root@fe222273db8e:/# hadoop version Hadoop 3.2.1 Source code repository https://gitbox.apache.org/repos/asf/hadoop.git -r b3cbbb467e22ea829b3808f4b7b01d07e Compiled by rohithsharmaks on 2019-09-10715:562 Compiled with protoc 2.5.0 From source with checksum 776eaf9eee9ceffc379bcbc1988737 This command was run using /opt/hadoop-3.2.1/share/hadoop/common/hadoop-common-3.2.1.jar root@fe2222773db8e:/# echo "Hello Hadoop Hello World" > sample.txt root@fe2222773db8e:/# dfs dfs -mkdir /input root@fe2222773db8e:/# hdfs dfs -mkdir /input 2025-09-01 11:58:14,080 INFO sa5l.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = remoteHostTrusted = false root@fe222273db8e:/# hdfs dfs -put sample.txt /input 2025-09-01 11:58:14,080 INFO sa5l.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = remoteHostTrusted = false root@fe222273db8e:/# hdfs dfs -cat /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-*.jar wordcot ut /output JAR does not exist or is not a normal file: /opt/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-* root@fe222273db8e:/# hdfs dfs -cat /output/part-r-00000 cat: `/output/part-r-00000': No such file or directory root@fe222273db8e:/# 9-	fals unt /
Output Analysis	Successful installation displays Hadoop daemons like NameNode DataNode, ResourceManager, and NodeManager running, and the web UI confirms cluster health and status.	
Link of student Github profile where	TO OF COMMING CLUSTER HOUSE HOUSE AND STATES.	





Yeshwantrao Chavan College of Engineering





NAAC A++ Ph.: 07104-237919, 234623, 329249, 329250 Fax: 07104-232376, Website: <u>www.ycce.edu</u>

Department of Artificial Intelligence & Data Science

Vision of the Department

To be a well-known centre for pursuing computer education through innovative pedagogy, value-based education and industry collaboration.

Mission of the Department

To establish learning ambience for ushering in computer engineering professionals in core and multidisciplinary area by developing Problem-solving skills through emerging technologies.

lab assignment has been uploaded			
Conclusion	Installation of Apache Hadoop on Linux System implemented successfully.		
Plag Report (Similarity index < 12%)	Apache Hadoop is a free software framework for distributed storage and processing of large sets of data using the MapReduce programming model. It works on a cluster of commodity hardware with a high degree of scalability and fault-tolerance. The Hadoop ecosystem primarily consists of HDFS (Hadoop Distributed File System) for distributed storage and YARN (Yet Another Resource Negotiated) for resource management and job scheduling. To install Hadoop on a Linux system you have to install Java and enable SSH and also modify the core configuration files for pseudo-distributed mode or fully distributed mode.	Unique 100% Exact 0% Partial 0% Congratulations	
Date	24 /07 / 25		