Nagar Yuwak Shikshan Sanstha's
Yeshwantrao Chavan College of Engineering
(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)
Hingna Road, Wanadongri, Nagpur - 441 110
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.

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 pronounce as Pep-si-lL easy to recall**
PEO2	**Core Competence**	**E: Environment (Learning Environment)**	
PEO3	**Breadth**	**P: Professionalism**	
PEO4	**Professionalism**	**C: Core Competence**	
PEO5	**Learning Environment**	**L: Breadth (Learning in 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)

Nagar Yuwak Shikshan Sanstha's
Yeshwantrao Chavan College of Engineering
(An Autonomous Institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University)
Hingna Road, Wanadongri, Nagpur - 441 110
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.

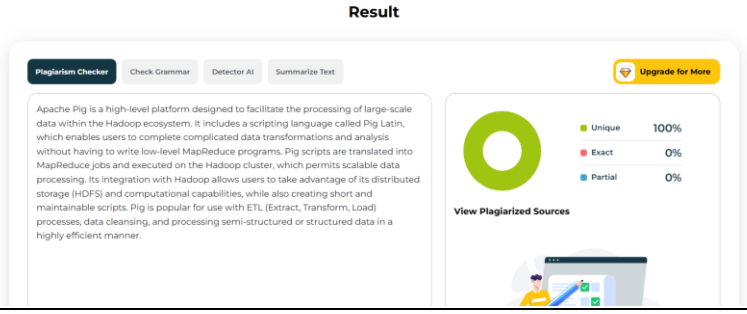
Session	2025-26 (ODD)	**Course Name**	BIG DATA AND HADOOP-LAB
Semester	7 AIDS	**Course Code**	22ADS704
Roll No	03	**Name of Student**	**Debasrita Chattopadhyay**

Practical Number	07
Course Outcome	1. Understand big data analytics and its business applications. 2. Analyze the HADOOP and Map Reduce technologies associated with big data analytics. 3. Apply Big Data analytics Using Pig and Hive.
Aim	Installation of Apache Pig on Linux with Hadoop Integration
Problem Definition	Installation of Apache Pig on Linux with Hadoop Integration
Theory (100 words)	Apache Pig is a high-level platform designed to facilitate the processing of large-scale data within the Hadoop ecosystem. It includes a scripting language called Pig Latin, which enables users to complete complicated data transformations and analysis without having to write low-level MapReduce programs. Pig scripts are translated into MapReduce jobs and executed on the Hadoop cluster, which permits scalable data processing. Its integration with Hadoop allows users to take advantage of its distributed storage (HDFS) and computational capabilities, while also creating short and maintainable scripts. Pig is popular for use with ETL (Extract, Transform, Load) processes, data cleansing, and processing semi-structured or structured data in a highly efficient manner.
Procedure and Execution (100 Words)	Steps of Implementation: - 1. **Prerequisites**: Hadoop installed and running; Java JDK installed with JAVA_HOME set. 2. **Download Pig**: wget https://downloads.apache.org/pig/pig-0.20.2/pig-0.20.2.tar.gz 3. **Extract Pig**: tar -xvzf pig-0.20.2.tar.gz 4. **Move to directory** (optional): sudo mv pig-0.20.2 /usr/local/pig 5. **Set environment variables** in ~/.bashrc: export PIG_HOME=/usr/local/pig export PATH=$PATH:$PIG_HOME/bin export HADOOP_HOME=/usr/local/hadoop 6. **Apply changes**:

	source ~/.bashrc 7. **Verify Pig installation**: pig -version 8. **Run Pig**: o Local mode: pig -x local o Hadoop mode: pig -x mapreduce 9. **Execute sample Pig script** (sample.pig): A = LOAD 'input.txt' USING PigStorage(',') AS (name:chararray, age:int); B = FILTER A BY age > 25; DUMP B;
	Code: # Go to /usr/local cd /usr/local # Download Pig 0.17.0 sudo wget https://downloads.apache.org/pig/pig-0.17.0/pig-0.17.0.tar.gz # Extract the tar file sudo tar -xvzf pig-0.17.0.tar.gz # Rename folder for simplicity sudo mv pig-0.17.0 pig # Set PIG_HOME echo "export PIG_HOME=/usr/local/pig" >> ~/.bashrc # Add Pig bin to PATH echo "export PATH=\$PATH:\$PIG_HOME/bin" >> ~/.bashrc # Set HADOOP_HOME (if Hadoop installed) echo "export HADOOP_HOME=/usr/local/hadoop" >> ~/.bashrc # Reload bashrc source ~/.bashrc # Verify Pig installation pig -version cd ~ echo -e "Alice,30\nBob,22\nCharlie,28" > input.txt

```
# Optional: verify content
cat input.txt

pig -x local

NOTE: At the grunt> prompt, type:
-- Load the input file
A = LOAD '/home/theia/input.txt' USING PigStorage(',') AS
(name:chararray, age:int);

-- Filter records where age > 25
B = FILTER A BY age > 25;

-- Display the results
DUMP B;
```

Output:

	
Output Analysis	**Interpretation**: Only people older than 25 are displayed. **Pig Features Used**: LOAD, FILTER, DUMP
Link of student Github profile where lab assignment has been uploaded	
Conclusion	Installation of Apache Pig on Linux with Hadoop Integration implemented successfully.
Plag Report (Similarity index < 12%)	
Date	**4 /9 / 25**