



Department of Computer Technology

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: To harness the power of artificial intelligence and data science to solve real-world problems and enhance human potential.	Mission: To acquire skills through coursework, projects, and internships, while actively engaging in research and collaboration with peers to innovate and apply AI solutions.
---	---

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.

Prerana Bijekar 30 October 2025

Name and Signature of Student and Date

(Signature and Date in Handwritten)



Department of Computer Technology

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	BDH Lab
Semester	7	Course Code	22ADS704
Roll No	11	Name of Student	Prerana Bijekar

Practical Number	4
Course Outcome	CO1: Understand big data analytics and its business applications. CO2: Analyze the HADOOP and Map Reduce technologies associated with big data analytics. CO3: Apply Big Data Analytics Using Pig and Hive.
Aim	Installation of Apache Hive on Linux with Hadoop Integration.
Theory (100 words)	Apache Hive is a data warehouse system built on top of Hadoop that provides a SQL-like interface (HiveQL) for querying and managing large datasets stored in HDFS. It converts SQL queries into MapReduce or Tez/Spark jobs, enabling users to perform data analysis without complex programming. Installing Hive on a Linux system integrated with Hadoop involves configuring environment variables, setting up Hive metastore, and connecting it to Hadoop's HDFS. This setup allows users to manage structured data efficiently using queries while leveraging Hadoop's distributed processing power.
Procedure and Execution (100 Words)	<p>Steps of implementation:</p> <ul style="list-style-type: none"> • Ensure Hadoop is installed and running. • Download and extract Apache Hive. • Configure environment variables in .bashrc. • Edit hive-site.xml to set metastore and warehouse paths. • Initialize the Hive metastore database. • Start Hive using the hive command. • Create and query tables to verify integration with Hadoop. <p>Code:</p> <pre>hadoop@phoenixNAP:~\$ hadoop version Hadoop 3.4.0 Source code repository git@github.com:apache/hadoop.git -r bd8b77f398f626bb77917 83192ee7a5dfaeecc760 Compiled by root on 2024-03-04T06:35Z Compiled on platform linux-x86_64 Compiled with protoc 3.21.12 From source with checksum f7fe694a3613358b38812ae9c31114e This command was run using /home/hadoop/hadoop-3.4.0/share/hadoop/common/hadoop-common-3.4.0.jar</pre>



Department of Computer Technology

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.

The figure consists of four vertically stacked screenshots illustrating the download process for Apache Hive:

- Screenshot 1: Apache Hive Downloads Page**
Shows the "Downloads" section of the Apache Hive website. A red arrow points to the "Download a release now!" button.
- Screenshot 2: Apache Software Foundation Downloads Page**
Shows the "hive" directory under "HTTP". A red arrow points to the "hive-4.0.0/" link.
- Screenshot 3: Index of /hive**
Shows a file listing for the "/hive" directory. A red arrow points to the "hive-4.0.0/" link.
- Screenshot 4: Index of /hive/hive-4.0.0**
Shows a file listing for the "/hive/hive-4.0.0" directory. A red arrow points to the "apache-hive-4.0.0-bin.tar.gz" file.



Department of Computer Technology

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.

	<p>Output:</p> <pre>hadoop@phoenixNAP:~\$ wget https://downloads.apache.org/hive/hive-4.0.0/apache-hive-4.0.0-bin.tar.gz --2024-09-02 07:57:53-- https://downloads.apache.org/hive/hive-4.0.0/apache-hive-4.0.0-bin.tar.gz Resolving downloads.apache.org (downloads.apache.org)... 88.99.208.237, 135.181.214.104, 2a01:af8:10a:39da::2, ... Connecting to downloads.apache.org (downloads.apache.org) 88.99.208.237 :443... connected. HTTP request sent, awaiting response... 200 OK Length: 458782861 (438M) [application/x-gzip] Saving to: 'apache-hive-4.0.0-bin.tar.gz' apache-hive-4.0.0-b 100%[=====] 437.53M 17.6MB/s in 26s 2024-09-02 07:58:19 (16.9 MB/s) - 'apache-hive-4.0.0-bin.tar.gz' saved [458782861 1/458782861]</pre> <pre>hadoop@phoenixNAP:~\$ tar xzf apache-hive-4.0.0-bin.tar.gz hadoop@phoenixNAP:~\$ ls -l grep hive drwxrwxr-x 11 hadoop hadoop 4096 Sep 2 08:00 apache-hive-4.0.0-bin → -rw-rw-r-- 1 hadoop hadoop 458782861 Mar 25 13:58 apache-hive-4.0.0-bin.tar.gz</pre>
Output Analysis	After successful installation, Hive starts with the prompt <code>hive></code> . Commands like <code>CREATE TABLE</code> , <code>LOAD DATA</code> , and <code>SELECT</code> execute successfully using HDFS as storage. Query outputs confirm proper communication between Hive and Hadoop, ensuring seamless data processing and retrieval.
Github Link	https://github.com/Prerana-Bijekar/BDH
Conclusion	Installing Apache Hive on Linux with Hadoop integration enables efficient management and querying of large datasets through a familiar SQL-like interface. This setup simplifies big data analytics by combining Hive's ease of use with Hadoop's scalability and distributed processing power.
Plag Report (Similarity index < 12%)	 <p>SmallSEOTools</p> <p>Plagiarism Scan Report By SmallSEOTools</p> <p>Report Generated on: Oct 31, 2024</p> <p>8.4% Plagiarized Content</p> <p>5.3% Exact Plagiarized</p> <p>3.1% Partial Plagiarized</p> <p>91.6% Unique Content</p> <p>Total Words: 388 Total Characters: 453 Plagiarized Sentences: 13 Unique Sentences: 113 (91.6%)</p>
Date	30 October 2025