

Innovative Teaching Practices

Faculty Name : Mrs.E.Amrutha Varshini
Course Name : Big Data Analytics
Class : III B. Tech I Semester
Academic Year : 2021-2022
Title of the Topic : Hadoop input and output
Activity Name : Collaborative Learning

Objective:

To create a collaborative and supportive learning environment where students work together to explore and solve problems related to Hadoop's input and output operations in Big Data Analytics (BDA). This approach will foster critical thinking, teamwork, and help students learn from each other's insights and experiences, enhancing their understanding of Hadoop's data processing mechanisms.

Steps to Implement Collaborative Learning for "Hadoop Input and Output in BDA"

1. Assign Problems

- **Divide the class into small groups**, each focusing on a different aspect of Hadoop input and output, such as data formats, input/output formats in Hadoop, and the integration of Hadoop with various data sources (e.g., HDFS, local files, and databases).
- **Provide each group with a specific problem** to explore. The problems will encourage students to analyze the input/output operations within Hadoop and solve real-world challenges using the Hadoop ecosystem.

2. Problems for Students Related to Hadoop Input and Output:

1. Hadoop Input Formats

○ **Problem:**

Explain and demonstrate how Hadoop handles input from different sources like local files and HDFS using various input formats.

Task: Each group should explore different **input formats** in Hadoop, such as TextInputFormat, KeyValueTextInputFormat, and SequenceFileInputFormat.

Question: How do these input formats affect performance and data processing in Hadoop jobs?

2. Hadoop Output Formats

- **Problem:**
Explore how data is written out from a MapReduce job in Hadoop.
Task: Each group should analyze **output formats** like `TextOutputFormat`, `SequenceFileOutputFormat`, and `MultipleOutputs`.
Question: What are the use cases for each of these output formats, and how can they be customized to handle large-scale output data efficiently?
- 3. **Custom Input/Output Formats in Hadoop**
 - **Problem:**
Design a custom input/output format to process a specific type of data, such as JSON or Avro files, in Hadoop.
Task: Implement a custom **InputFormat** or **OutputFormat** for a specific data type.
Question: How does creating custom formats improve the flexibility and performance of Hadoop for diverse datasets?
- 4. **Integration of Hadoop with Other Data Sources**
 - **Problem:**
Discuss how Hadoop can be integrated with external data sources such as **HBase** or **Relational Databases** for input/output operations.
Task: Demonstrate how to configure Hadoop to read from and write to an **HBase table** using the appropriate input and output formats.
Question: What are the advantages and challenges of integrating Hadoop with external data sources for input and output?

3. Collaborative Activity

- **Group Work**
Each group will collaborate to explore their assigned problem. They will:
 - **Research the input/output formats** and the mechanisms behind Hadoop's data processing.
 - **Experiment with code samples** to demonstrate input/output handling for specific data sources.
 - **Present solutions** with clear examples of how different formats are used in real-world Hadoop jobs.
- **Documentation:**

After working through the problems, groups will document their findings and prepare a brief report or slide deck summarizing their approach, solutions, and key insights.

4. Discussion and Presentations

- **Group Presentations**
After completing their problems, each group will present their findings to the class, including:
 - An explanation of the input/output formats they researched.
 - Code snippets or examples that demonstrate how to handle input/output in Hadoop.
 - Insights into the benefits and challenges of using different formats for big data processing.
- **Class Discussion:**
Encourage active discussions after each presentation, where students can ask questions,

offer suggestions, and share additional experiences related to Hadoop input/output operations.

5. Feedback and Reflection

- **Constructive Feedback**

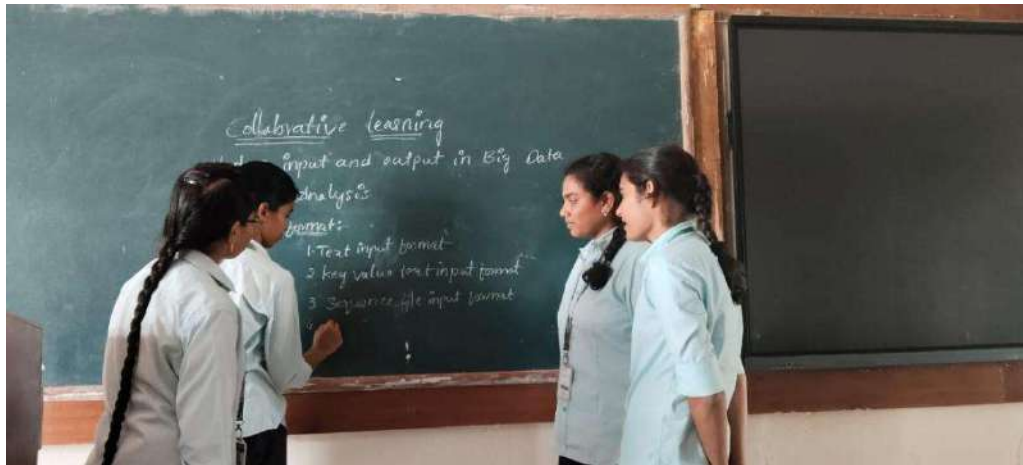
Provide feedback on the approaches taken by each group, focusing on the accuracy of the concepts and the practical application of Hadoop input/output operations.

- Offer suggestions on how to improve data processing workflows in Hadoop based on their findings.

- **Reflection**

Reflect on how various Hadoop input/output formats can be used to address real-world data challenges, improve performance, and optimize big data workflows. Encourage students to think about scenarios where different formats and configurations would be ideal for specific types of data.

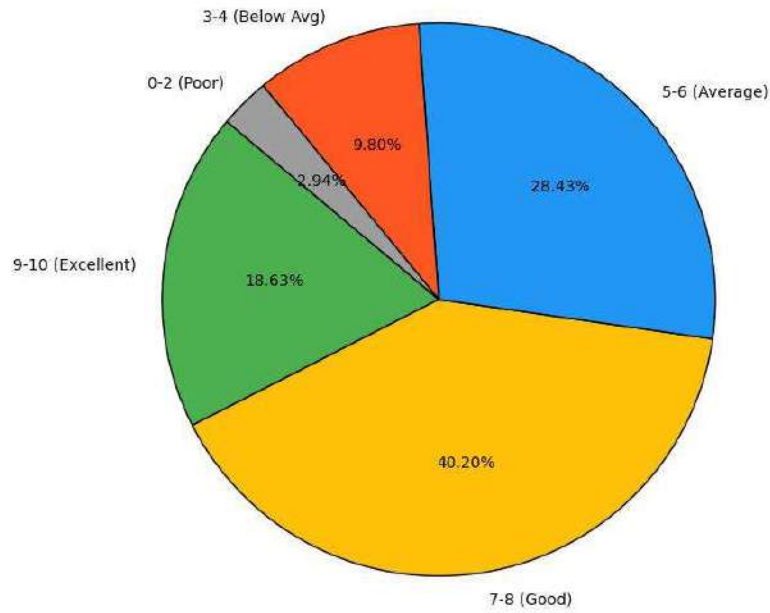
Screenshot of the practice



Assessment Summary

Marks Range	Number of Students	Percentage
9-10 (Excellent)	19	18.63%
7-8 (Good)	41	40.20%
5-6 (Average)	29	28.43%
3-4 (Below Avg)	10	9.80%
0-2 (Poor)	3	2.94%
Total	102	100%

Marks Distribution Among Students



Conclusion:

This collaborative learning approach will help students gain a deeper understanding of **Hadoop's input and output mechanisms** in Big Data Analytics. By working together to explore different input/output formats, designing custom solutions, and presenting their findings, students will enhance their problem-solving and teamwork skills while mastering crucial concepts for handling large-scale data processing in Hadoop.

Signature of the Faculty

Head of the Department