



# An Integrated Way for Teaching Hadoop & BigData Analytics Course

E. Laxmi Lydia, N. Sharmili, T. V. Madhusudhanarao, Madhu BabuChevuru, K. Vijaya Kumar

**Abstract:** Hadoop & BigData Analytics course has occupied a ubiquitous role in present software technologies. Educational institutes are fond of this course as it's been trending course most of the placements in software companies are based on. As per the traditional teaching mechanisms, the educational systems are not much up-to-mark where the students are not assisted with the course resulting in atrocious placements. Therefore to enhance this placements, institutions has to adopt a new integrated teaching- learning proceedings which help in drastic change of academic results discussed in this paper. Here the result analysis of course attainments are compared to show the eye-catching improvements as occurred in VIIT College.

**Keywords:** Hadoop, Big Data, Analytics, technologies, academic, attainments

## I. INTRODUCTION

In most of the educational institutes, traditional teaching practices are been carried that impacts the result analysis along with practical knowledge of students. Although the institutes appoint well knowledge teachers, they can't concentrate on each of the students in the class whole day. They can maximum concentrate on 10 to 15 students in the class. We know that the learning capabilities of students vary in class irrespective of same teaching practice followed by the teachers. This shows that teacher as to spend time on slow seekers when compared to fast learners which leads to disturbances in syllabus completion or makes class boring to others. Therefore leading to instinct downfall. To avoid these fluctuations, we have approached you with four innovative teaching learning mechanisms.

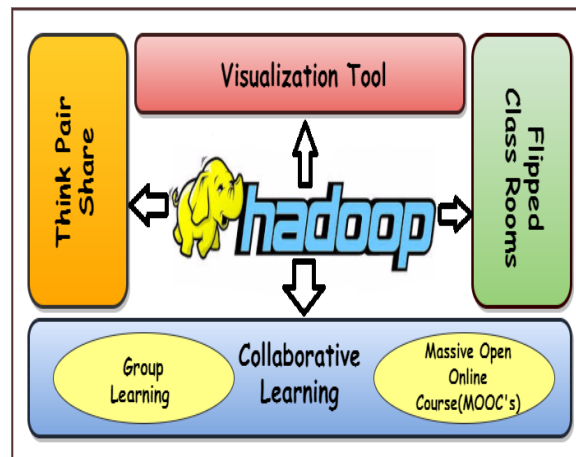


Figure. 1: Innovative Teaching-Learning Mechanisms

Figure 1 shows the different innovative teaching-learning mechanisms that are discussed in this paper. They are: Think Pair Share, Flipped Classrooms, Visualization Tool and Collaborative Learning. These mechanisms are been well fledged used in college VIIT for Hadoop & Big Data Analytics course. The outcomes of these mechanisms are shown in figure 7 with course attainments outcomes achieved in the last two years. Table 1 demonstrating the teaching-learning modes with the detailed role of teacher, student and outcome of the learning activity.

### Mechanism 1: Think Pair Share

Think Pair Share mechanism helps students to elevate themselves with new approaches when compared with traditional teaching techniques. This makes student to understand critical topics easily with discussions held in the classroom. According to this mechanism, a troubleshooting question is been given by the teacher to the students. He asks them to resolve the question and find out the solution individually. After solving the question, he asks each student to compare their results with neighboring student (making a pair) and discuss their approaches to find the appropriate solution. Lastly, one of the pairs in class come-up with a solution and explains it to whole class.

Revised Manuscript Received on 30 July 2019.

\* Correspondence Author

**Dr. E Laxmi Lydia**, Associate Professor in CSE, Department, Vignans Institute of Information Technology, Visakhapatnam.

**Dr. N. Sharmili**, Associate Professor in CSE Department, Gayatri Vidya Parishad College of Engineering for Women, Viskahpatnam, Andhra Pradesh, India.



**Dr. T V Madhusudhanarao**, Professor in CSE Department, Vignans Institute of Information Technology, Visakhapatnam.


**Madhu Babu Chevuru**, Asst. Professor, Department of Computer Science Engineering, VFSTR deemed to be University.

**K. Vijaya Kumar**, Associate Professor, Department of Computer Science Engineering, Vignans Institute of Information Technology for Women, Andhra Pradesh, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Table 1 teaching-learning mechanism

Phase	Problem	Time Duration (in min)	Teacher Role	Student Role	Outcome
Think Phase 	X, Y, Z are the hospitals that are well sophisticated. They maintain the records of each patient detail in computerized format. The patients mostly admitted are related to cancer prone. They are advised to go for weekly check-up for the status of their chronic. They are given chemotherapy based on their chronic status. These chemos are given in alternate hospital based on their working hours. For these they have to maintain a huge data of each patient. And also shareable to other hospitals for further process. Here how can the doctors know about their patient details quickly based on his previous status? A) Graph B) Data Analysis C) Graph and Data Analysis	10 min	Problem statement given. Ask students to resolve the problem individually .	Within given time the student individually have to find a solution and put it on paper.	Best representation of previous patients status quickly
Pair Phase 	Here the individual student has to verify his solution with his neighbor for correct answer. This way every student make a pair with his neighbor and resolve the solution if the obtained outputs are not same. They have to make a one final output.	10 min	Makes student to verify /compare their answers with their neighbor. If the answers are same they can conclude with the output. If not same then they have to discuss, analyze and come out with final output.	If same answer can conclude with result obtained. Otherwise the pair has to discuss, rethink and analyze the problem with their neighbor.	One finalized solution is agreed with the pair.

<p>Share Phase</p> 	<p>Few pairs are called on to stage and explain their output in front of all for about a minute each. In this way student pair shares their answers with all.</p>	<p>15 min</p>	<p>Teacher has to make a note of all the student outcomes and have to discuss it at the end of the session. He has to explain the correct output and make clear the students of their mistake in their solutions.</p>	<p>The pairs have to check their results with other pairs when explained and note down their faults.</p>	<p>Correct solution (C) is obtained at end of the session.</p>
--	---	---------------	---	--	--

Here the students are motivated to think as a part of activity. Here as per the problem the students get an idea about the BigData analytics subject that both data and graph are important for quick analysis.

**Mechanism 2: Flipped Classroom**

Flipped Classroom Mechanism is awareness activity that makes students have knowledge based on outside and inside class activities. So that students get a clear and practical knowledge of subject.

**Outside Class Activity:**

Uploaded Video related to Course:  
<http://www.youtube.com/watch?v=thiTsWC1bnU>

**Key Concepts:**

- Data representation of the given input.
- Algorithms suited for the analysis of that particular data analysis.
- Implementation of correct concepts.
- Have a clear idea of database and data mining concepts
- Programming concepts of MapReduce.
- Weather report example, Hospital patient data, stock market data analysis.

**Inside Class Activity:**

- Should explain the Hadoop concepts in detail.
- Hadoop Ecosystem concepts.
- MapReduce programming concepts.

**Activity Plan:**

- <http://www.youtube.com/watch?v=thiTsWC1bnU> : Students are advised to come to next class by watching the above video. By watching the above video student will get a basic idea related to the topic to be discussed in next class.
- Trouble shooting question is posed on students to solve in 5 min by the teacher.

- Students are asked to resolve the solution individually.
- After completion of individual analysis, they are asked to discuss their answers with their neighbor for 10 min.
- Few pairs are asked to share their results in front of all students in class.
- Later the teacher explains the finalized output by rectifying the faults made by students.

**Mechanism 3: Visualization Tool**

Visualization tools help the students to gain the concepts quickly when compared to others; some of the visualization has been represented in the figure 2,3,4,5 and 6. As the students are well attracted to the diagrammatic view of the concepts, the subject will be rooted strongly in brain. Here are some screenshots of the tool showing MapReduce, Hive, and DFS.

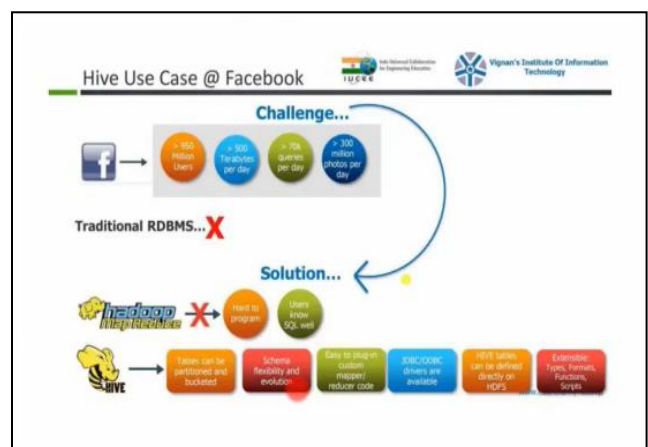


Figure 2: Visual representation of hive introduction using usecase

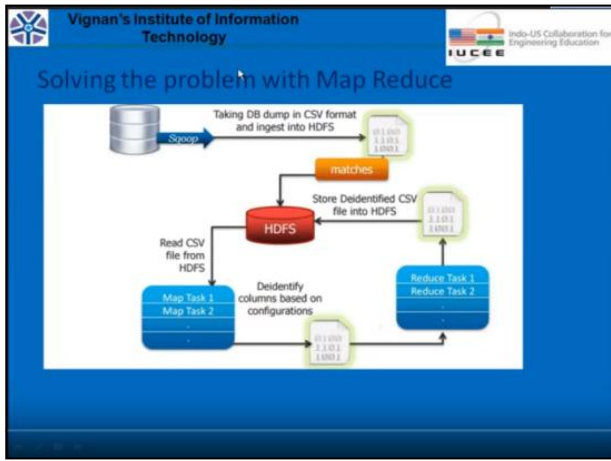


Figure 3: Visual representation of architectural design for HDFS flow

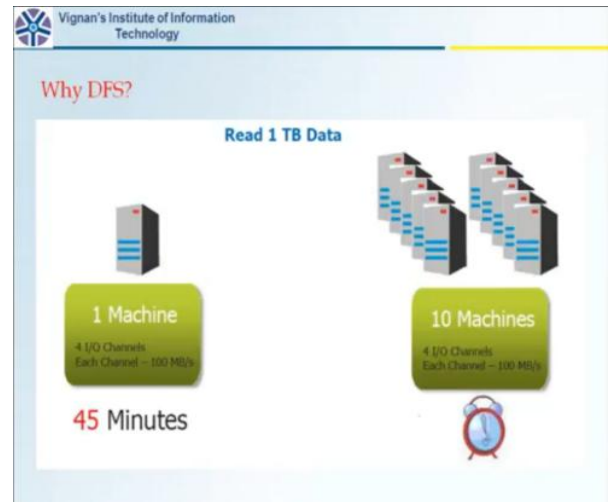


Figure 6: Visual representation of Distributed File System



Figure 4: Visual representation of cloudera

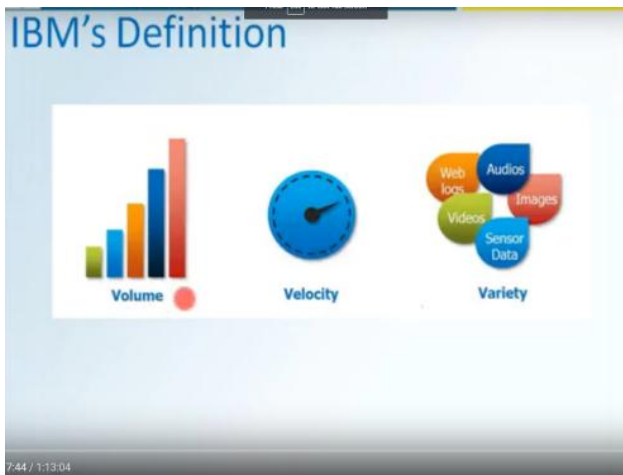


Figure 5: Visual explanation for IBM's definition

**Mechanism 4: Collaborative Learning**

Collaborative Learning is a practical session for the students in the Hadoop & BigData lab to enrich the programming skills of the students.

**a) Group Learning**

- In the lab, divide the student into 5 groups (having 3 to 4 members in each group) based on the number in class.
- 5 problems are given to these 5 groups randomly and asked them to solve each problem.
- The students start discussions related to the given problem and find out a best algorithm that suits.
- After resolving the algorithm, students will individually implement the program on their own systems.
- One of the groups who complete their task successfully would explain the algorithm and program to all other students.
- The explanation of the problem and solution is shared among the members of the group, in such a way that all group members participate and share their ideas with others.
- Through this activity the weaker students also get a chance to mingle with other students and can gain knowledge in programming skills.

**b) Massive Open Online Course (MOOCs)**

MOOCs is a model for delivering learning content through online to any person with n limit on attendance. VIIT is JNTU Kakinada affiliated college, which organized MOOCs for Hadoop & BigData course during the academic year 2016-2017 and 2017-2018 which emphasized on developing the ability to write programs in order to solve the computational problems. The MOOCs course made the student to write the several programs on their own. This course deals with several programming paradigms such as MapReduce, Pig, Hive. The course offered as MOOCs helped the students to improve programming skills as the training given by the industry experts which as removed the gap between industry and the Academia.

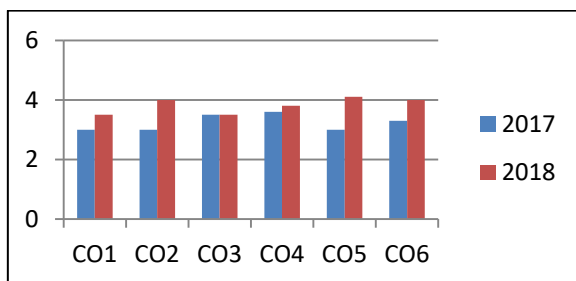
**II. RESULT ANALYSIS**

In our Vignan Institute of Information Technology we have designed special software to measure the course outcomes, Program Education attainments. The main aim is to satisfy the stake holder names as per the Outcome Base Education (OBC) standards. To achieve the defined outcomes which are demonstrated in table1, we are concentrating on various Teaching-Learning Practices at our Vignan Institute of Information Technology to improve the student’s knowledge. To quantum the impact of innovative Teaching-Learning mechanisms various comparisons have been made for the course attainments directly and indirectly. The result analysis of last academic years i.e., 2015-16, 2016-17 and 2017-18 is carried out for the Hadoop&BigData course to measure the student understanding with respect to course attainments. Analysis has been made on total 240 students. Table No.2 represents the predefined course outcomes which have to be satisfied by the direct stake holders.

**Table No.2: Predefined Course outcomes**

	Course Outcomes
CO1	Develop Data Structure concepts through JAVA
CO2	Analyze the configuring of Hadoop clusters effectively
CO3	Illustrate Hadoop API for MapReduce framework
CO4	Develop basic programs of MapReduce framework particularly driver code, mapper code, reducer code
CO5	Building a complete business data analytic solution and apply structure of Hadoop data with Hive

To measure the innovative teaching learning mechanisms, we took the previous year course attainments both directly and indirectly. We also took survey from the students based on course taught.



**Figure7: Course Outcome Attainments**

From Figure 7 we can get an idea of course outcomes analysis in last two years. CO1 deals with real world concepts that relate to BigData where we can see 0.5 increases in 2018 compared to 2017. CO2 deals with development of data structure concepts through Java and CO3 deals with analysis of configuring of Hadoop clusters effectively. These CO’s relate to collaborative mechanism

where we can see there is improvement of 1 point. CO4 and CO5 deals with MOOCs concept during lab sessions where we can see enhancement in 2018. CO6 deals with whole practical session related to programming along with graph analysis through Think pair phase and Flipped classrooms. As when we compare with program specific outcomes (PSO) we gain improvements when compared to previous year 2017.

**III. CONCLUSION:**

Old teaching methods have many drawbacks and deficiencies. In this paper novel mechanisms have introduced that enhance the educational institute placement results along with students’ knowledge. With present modern culture and developments, these innovative ways will help teachers to activate each of the students in class and to interact with each of them. We can also see the improvement in academic status outcomes by following these mechanisms (Think Pair Share, Visualization tools, Flipped Classrooms and Collaborative learning). Of these Visualization tools and MOOCs helped a lot to students in programming skills development.

**REFERENCES**

1. Abdurrahman Hi, “Using the Think Pair Share strategy to improve students’ speaking ability at stain ternate”, Journal of Education and Practice, ISSN 2222-1735 (PAPER) ISSN 2222-288X, Vol. No 6.
2. Webb N.M., “Peer interaction and learning in small groups”, International Journal of Education Research, 13 (1), Page no: 21-39.
3. C.Demetry, Work in progress: An innovative merging “classroom flips” and tea-based learning. In Proceed-ings, 40<sup>th</sup> ASEE/IEEE Frontiers in education Conference,2010.
4. N. Warter, Perez and J. Dong, “Flipping the classroom: How to embed inquiry and design projects into a digital engineering lecture”, in proceedings of the 2012 ASEE PSW section conference, 2012.
5. V. T. Lokare and P. M. Jadhav, “Course outcomes attainment for data structure course using direct and indirect methods”, in Journal of Engineering Education Transformation, 2016.

**WEB REFERENCES (did by author):**

- <https://www.youtube.com/watch?v=10SWQfJUK0g>
- <https://www.youtube.com/watch?v=b1paiDb3Zd0>
- <https://www.youtube.com/watch?v=thiTsWC1bnU>
- <https://www.youtube.com/watch?v=2MsvFIA98LA>

