

# A Framework for Managing and Analyzing Big Data in Indian School Education System with Reference to Jammu & Kashmir

Jahangir Kamal, Meenu Dave

**Abstract:** Big Data storm has reached all most in every sector whether be public or private and is an important decision making factor for the administration or governing body of any system or organization, because with the advancements in technologies, various public, private, and social organizations are creating or exchanging a huge volume of data through different sources in various formats. Big Data can be applied in various sectors of India, wherein one of the essential sectors is education system where Big Data analytics is slowly and steadily finding its place for betterment of the services being provided by this system. India has one of the biggest school systems in the world which is spread over different states of the country and one part of this system is functional in one of the northern states of India known as Jammu and Kashmir (J&K). This study is based on Big Data management and analysis in School Education System of J&K. In the area of School Education in J&K, computers are finding an important place for data management as well as imparting learning through digital means. There is a need of applying Big Data technologies to various aspects of the School Education in J&K, as huge amount of data with variety and high frequency of generation is available in the institutes under this system. This paper will analyze various aspects of Big Data management and analysis system for School Education of J&K. This paper will also highlight the current scenarios of School Education System of J&K in handling Big Data such as sources of Big Data, Scope of Big Data analytics in School Education System of J&K, opportunities of Big Data analytics, challenges and issues that can be faced in this system, Applications of Big Data analytics in School Education of J&K, and finally discussing the proposed architecture of Big Data management and analysis system for School Education of J&K.

**Index Terms:** Big Data, School Education, Analytics, Hadoop, HBase.

## I. INTRODUCTION

Presently Big Data and its associated technologies have emerged as the basic power to drive the business and other sectors successfully. Various organizations are using Big Data as an information tool to analyze data and extract the effective information for efficient calculations, planning, and decision making, thus Big Data analytics has now become a famous term in almost every sector or organization. Big Data can be described as the generation of data in huge volumes like Gigabytes, Terabytes, Petabytes, and so on, with high speed or frequency from various different sources in different varieties like structured, semi-structured, and unstructured

types, which cannot be collected, stored, managed, and processed with conventional database software systems. Some organizations collect and manage Big Data [1] and provide interfaces to everyone to access and analyze some of the data, e.g. Google Insights provided by Google freely, while as some organizations do not provide any access to data. With advancements in Big Data technologies, issues and challenges arising during the phases of collection, management, and analysis of huge volume of data with variety and high velocity of generation are being effectively dealt with. Big Data has brought huge changes to the working standards of industries in the world.

There are various sectors in India where Big Data analytics can be applied for their all-around development and improvement in functioning such as Healthcare, Agriculture, Electricity, Education, and so on, as shown in fig.1. In comparison to other sectors in India, Big Data adoption and operations in education sector is not so prominent and is slow in action for the success of this highly essential service of the nation. Education sector in India is also aiming to make effective use of Big Data technologies for managing and analyzing the large amount of data coming out from different education systems like school education, colleges, and other higher educational institutes. Big Data analytics can be used to solve issues of education system [2], for example student retention problem in Georgia state university was solved through usage of data analytic tools and techniques. The data generated during the learning process of students can be used as an asset in modern School Education System of India. So far, Big Data management and analytics process has not been applied to that extent in the sector of School Education in India and thus not showing much successful impact on Indian School Education System.



Fig.1: Big Data in various Domains

Being one of the biggest school systems in the world, Indian School Education System is implemented in its different states where a separate administration or governing body is managing the school education system of respective states. Jammu and Kashmir (J&K) being one of the Indian

Revised Manuscript Received on June 15, 2019.

Jahangir Kamal, Computer Science/IT, Jagannath University, Jaipur, India.  
Dr. Meenu Dave, Computer Science/IT, Jagannath University, Jaipur, India.

states has two administrations for running the School Education System of the state, one is for Jammu region and another for Kashmir and Ladakh regions, which are known as Directorates of School Education in J&K. Table-I shows the total number of schools (primary, middle, high, and higher secondary schools) in J&K during 2017-18 with divisional distribution and table-II shows the district level distribution of these schools (Source: SMIS J&K). The condition of School Education in Jammu and Kashmir is horrifying to some extent with more and more students are getting enrolled in schools than ever before without proper structure of education system, lack of infrastructure in institutes, and less focus or attention given to the majority of students in rural areas, except some of the private schools. Table-III shows progressive enrollment of students at primary level in School Education of J&K (Source: Annual data collection under SSA & RMSA). The most literate state, Kerala, has highest literacy rate in India and is showing constant progress in their education system as compared to education sector of J&K. Table-IV shows total number of schools (Govt. as well as Private) in Kerala and J&K during the year 2007-08 [3].

Presently, School Education in J&K is becoming technology oriented to some extent where various schools mostly private institutes are using technology for handling some physical tasks in institutions through Learning Management System such as students are downloading and uploading their assignments, downloading of study material from institution web portal, uploading documents, quiz formation, attending exams online under the trained supervisors, and so on. Various institutes mostly private schools have adopted to teach students through multimedia devices and thus have changed blackboards to whiteboards and now to smart-boards. School libraries, student monitoring and attendance system has now been digitalized or automated in various institutions for the convenience of students and their security in the campus. These technology driven campuses of School Education in J&K work faster and smarter than traditional systems. The interaction of students with various working sections of institution like admission section, account section, course details, attendance, assignment submission, examination and results, library interaction, and so on; generate huge amounts of heterogeneous data and thus has transformed the modern school education into the Big Data era, which needs efficient and effective management and analysis processes in order to provide better services to the students and various benefits to the institution. Due to the high demand for transparency and accountability in every organization or system, School Education in J&K also aims to increase the same for better governance. School Education System of J&K can use Big Data and its associated technologies for measuring the efficiency of the system and can also make use of funds and facilities in fair manner. So, now there is a huge interest of using Big Data in School Education System of J&K in order to extract the value from the huge amounts of heterogeneous data to be used for betterment of the school education and the primary focus target users are students, teachers, and the School Education of J&K as a whole.

Kashmir	6097	4343	840	343	11623
Jammu	7363	3791	897	306	12357
Total	1346	8134	1737	649	23980

**Table-II: District Wise Count of Schools in J&K during 2017-18**

S. No.	District Name	Primary Schools	Middle Schools	High Schools	Higher Sec. Schools	Total
1	SRINAGAR	212	210	73	30	525
2	GANDERBAL	348	172	38	15	573
3	BANDIPORA	394	337	47	18	796
4	BUDGAM	710	432	90	41	1273
5	BARAMULLA	1128	648	162	44	1982
6	KUPWARA	993	713	93	45	1844
7	ANANTNAG	706	692	115	51	1564
8	PULWAMA	505	247	61	33	846
9	KULGAM	363	358	47	26	794
10	SHOPIAN	318	147	37	10	512
11	LEH	153	121	32	14	320
12	KARGIL	220	266	45	16	547
13	JAMMU	961	391	160	69	1581
14	UDHAMPUR	910	469	107	32	1518
15	KATHUA	912	390	107	37	1446
16	SAMBA	289	151	41	17	498
17	REASI	689	353	71	19	1132
18	RAJOURI	1013	560	112	41	1726
19	POONCH	940	458	70	24	1492
20	DODA	667	464	108	39	1278
21	KISHTWAR	428	303	58	15	804
22	RAMBAN	510	252	63	13	838

**Table III: Student enrollment in primary level of School Education of J&K**

Year	2015-16	2014-15	2013-14	2012-13	2011-12
Student admissions	1,373186	1,305169	1,223462	1,218785	1,239946

**Table-IV: Total no. of schools in Kerala and J&K during 2007-08**

State	Govt. schools	Local body schools	Private aided schools	Private unaided schools	Total
Kerala	5473	105	7865	1444	14887
J&K	16648	Nil	326	1652	18626

**Table-I: Total number of Schools in J&K during 2017-18**

Division	Primary Schools	Middle Schools	High Schools	Hr. Sec. Schools	Total

## II. SOURCES OF BIG DATA IN SCHOOL EDUCATION OF J&K

Big Data in school education means generation of

voluminous data with high variety from the learning process of students. Data existing in school educational institutes differ from the data generated from other organizations in terms of size, source, format, etc. Most of the data related to students and the institute is managed and stored manually on large stationary files and ledgers and thus lack the ability to harness the opportunities hidden within this huge amount of data through analysis and also this whole data remains vulnerable to various destructive attacks like fire, floods, etc., thus can waste time and money of the whole system of School Education in J&K. J&K being a conflict zone has suffered huge losses on part of educational infrastructures. In 1990, about 700 schools were damaged due to a sudden uprising and recently in 2016, J&K suffered a loss of 16000 crores due to an uprising and continuous shutdowns during which schools were closed for about 5 months and 31 school buildings were also burnt down [3]. In 2014 floods, various schools of J&K suffered huge damages and lost all the data of institutes including all the student records. Also through this manual data management in school education institutes a possibility

of misgovernance arises and is prone to fraudulent means. The institutes are making availability of various resources for the students like infrastructure facilities (classrooms, sitting arrangements, boards, etc.), human

resource (teaching and non-teaching staff, helpers, etc.), technical support (labs), course materials, etc. and in return, students provide an active participation and preparation for the smooth functioning of the institute in addition to the fees provided towards the finance of the institute. [3] Table-V shows the percentage of schools at different levels in Kerala and J&K with computer and library facilities. The infrastructure provided to students by the institutes should deliver better instructions, improve student outcomes and reduce drop out rate.

Table V: Percentage of schools in Kerala and J&K with computer and library facilities, 2014

Facilities Available	Kerala				J&K			
	Primary schools (%)	Upper Primary schools (%)	Secondary schools (%)	Higher secondary schools (%)	Primary schools (%)	Upper primary schools (%)	Secondary schools (%)	Higher secondary schools (%)
Computers	89.6	98.6	100	83.64	3.74	11.19	41.94	50
Libraries	94.69	98.93	98.20	96.36	42.27	54.48	61.29	100

Emerging online learning in education sector is creating huge digital data, which can be used for analysis and model formation. Due to the hybrid education system (online as well as on-campus learning) present in some institutes of School Education of J&K, analysis of education Big Data is possible at real-time to improve the quality of education. School education data sources [4] include two aspects like explicit behavior (e.g. test scores, homework completion status, classroom performance) and implicit behavior (e.g. forum posting, extra-curricular activities, online social networking activities) which is not directly considered in educational activities. Data generated through school education sector is classified into two categories such as static data and fluid data; [5] the data collected, recorded, and stored by educational institutes for future use is called static data which is further divided into following types:

- **Student record data:** This type of data is regarding the profile or personal data of students such as name, address, age, school, course, grades, contract details, attendance records, etc.
- **Staff data:** It involves the details about the number of working staff (teaching and non-teaching manpower), their biodata, subject or course with which they are affiliated, etc.
- **Admission and Application data:** It is related to number of applications for the institute in an academic year, number of students who could qualify for admission and those who acknowledged their acceptance, etc.

**ancial data:** This data is regarding the funds and finances such as funds provided, incomes, expenditures, budget, taxes, profits and losses in a particular year, etc.

- **Alumni data:** This data is about those students who have already completed a level of education from the institute also known as pass-out students. The data about them is usually their current contact details, name of enrolled institute, job positions if working, workplace, etc.
- **Course data:** It includes data about each course undertaken by the students and the number of students enrolled in each course.
- **Facilities data:** This data is about infrastructure facilities available in institutes such as number of classrooms and labs, equipments and softwares provided in the labs, etc.
- **Co-curricular data:** It involves data about co-curricular activities like number of events or activities taking place in the institute, participating students per activity, award winning students, etc.

Flow data in school education gets created when the students interact with the automated system or digital learning system of the institute and includes login details to the online course page, notes downloaded, biometric attendance, the time they stayed online for, the number of books issued in his name, etc. Presently in School Education System of J&K, most of the institutes are managing much more static data than the fluid data, because of lack of proper infrastructure and skilled manpower, for example Syllabi, course materials, learning

•  
F  
in

activities, curriculum, examination patterns, attendance, results, timetables, Administration data, staff data, facilities available, etc. School education system or any other organization should be aware of types of data with which to deal with such as structured data (database), semi-structured (CSV files), and unstructured data (text, images). Presently data is available in huge amounts in School Education System of Jammu and Kashmir that is having new characteristics like faster availability, greater scope, possessing new observations and with more complex structure, which differentiates it from the traditional time-series or panel data. This huge amount of data needs efficient management and analytical tools to harness the benefits for the School Education in J&K.

School Education of J&K generates variety of data which include structured, semi-structured, and unstructured data like database data, text, images, websites, etc. Unstructured and semi-structured data of school education in J&K is currently ignored by not managing and analyzing it for decision making, whereas it contains essential information like learning styles and learning habits of students. Most critical aspect of education data is decision time, [2] which is near real time in nature such as daily course attendance, online learning activity, campus behavior and so on. Real-time or near real-time data collection remains favorite for any system or organization so that timely analysis and decision making can be possible for the success. In School Education System of J&K, real-time data availability is much less than other sectors, because of many factors like manual system of data recording, non-automation of school management, no online facility for data management, and so on. Through analysis of collected data, future performance of students can be forecasted and also problems faced by students and their academic status can be timely and accurately assessed.

Education Big Data open service model as shown in fig.2 can be used in School Education System of J&K as it provides solutions to the data collection problem faced by staff and researchers or analysts. This platform was tried at the Central

### III. BIG DATA ANALYTICS IN SCHOOL EDUCATION OF J&K

“The goal is to turn data into information and information into insight”, (Carly Fiorina) [2]. Big Data is a huge volume of heterogeneous data from which useful information or value can be extracted. Big Data Analytics is defined as a process to have a thorough investigation of large data sets for identifying the hidden pattern and other necessary information [5], which can be used to make important decisions, thus helping in improving the system and saving important assets of the concerned sector. Two types of innovations are possible by using analytics which also separate systems from mainstream-one: is sustainable (improves the existing system) and the other is disruptive (creates new systems different from the existing system). The main steps followed in an ideal Big Data model are-*Strategic question*: The query around which the whole idea revolves, *Data analysis and Prediction*: Identifying and interpreting the data pattern and using the predictions over the same, *Insight and Action*: After a thorough evaluation and understanding of the identified pattern necessary and needful actions are taken, which is the implementing phase of analytics. The prediction and decision making in school education is possible through Big Data analytics process which can be defined for an education sector

China Norma university and offers assistance to the key aspects of education sector like teachers and students [2]. Suitable IT infrastructure is needed either at administrative or institution level to harness the benefits of Big Data management, processing, and analysis processes in School Education System of J&K with high skills of data preparation, as 90% of total time spent on project gets involved in data preparation tasks (data collection, integration, cleaning, etc.).

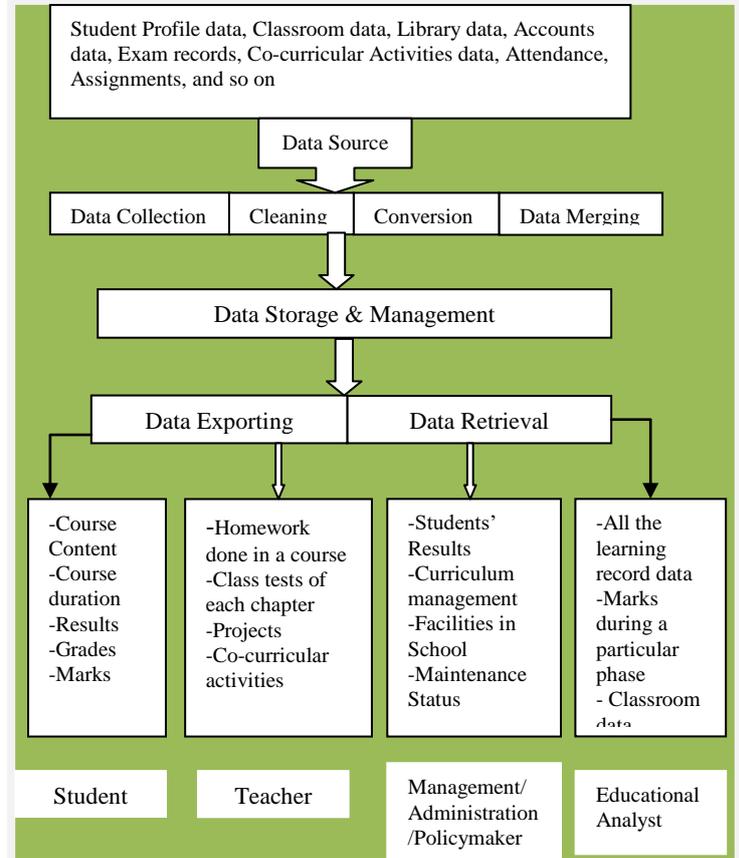


Fig.2: Education Big Data Open Service Model

as a process to examine large data sets, generated by students through interaction with the course, to identify the learning patterns and other necessary information important for decision making which can eventually boost the school education sector at a low cost, thus saving time and money of the system. Analytics can be used in school education system to identify the achievements and areas of weakness by comparing the results with other institutes and thus students, teachers, and institutes can check the progress of academics and student behavior by conquering Big Data storm. Big Data analytics is a broad concept embracing various processes which is also called as data-driven decision making, where data is collected, processed and then turned into information which is used by learners and teachers for betterment and smooth progress of learning.

There are six levels of analytics in education sector defined as follows [5]:

1. *Personal level*: Analysis carried out on personal performance according to the goals and labels set by an individual.
2. *Course level*: It involves the conceptual development.
3. *Departmental level*: It deals with predictive modeling and patterns of success or failure.
4. *Institutional level*: Analysis of performance of academics, resource allocation, etc.

5. *Regional level*: Analysis at this level compares systems and standards with each other.
6. *Specific levels*: Other different levels are defined more specifically as:
  - *Nanolevel*: Analysis of activities in a course.
  - *Microlevel*: It describes an entire course in the education program.
  - *Mesolevel*: Describes many courses in an academic year.
  - *Macrolevel*: Describes several programs in the institution.

Big Data is a powerful tool that has the ability to transform the school education in Jammu and Kashmir (J&K). The administration or governing body of School Education in J&K should pursue some objectives to harness the power of Big Data analytics and also need to understand the opportunities and importance of the data under their possession, so that they discover new ways to the next generation of analytics. By using Big Data analytics in School Education of J&K, a complete view of each and every activity at institution level as well as management or administration level is possible and thus all components of the school education system (students, faculty, and administration) can make better decisions at right time. Some proper suitable models and the necessary tools and techniques should be identified to establish a data driven decision making and forecasting. Some challenges or hurdles like applying analytics on student academics and ethical issues related to the use of data of students make execution of Big Data driven decision making difficult in School Education System of J&K. An efficient Big Data management and analysis system with some principle guidelines for institutes can help administration to overcome these challenges. Analysis of Big Data generated through School Education in J&K should be realized as a profitable investment in spite of a costly affair. Big Data analytics is an effective measure for School Education in J&K to achieve their goals, because the huge amount of data possessed by school education has a huge potential of impact and is an important factor to bring accountability and transparency in the governance of School Education System of J&K.

The main aim of analytics in school education is to help teachers and administration to generate opportunities in education for each and every student according to their abilities and needs. In the current era of Big Data, school education data analysis is operated around the collected data which needs to be complete in all aspects and all the information should be collected. School Education System of J&K should have a good understanding about Big Data and its applications, so that best strategies can be determined through efficient management and proper analysis, which is also time specific because some data requires frequent analysis for determining faster changes in the learning environment. Two kinds of approaches can be used in analytics process of school education data; one is totally data driven approach of analysis which begins with data and ends up on a decision reached, so data is important here and also the methods like collecting, cleaning, managing, and processing this data. Second kind of approach is need based approach or context based approach of analysis which is carried out when a need for a decision arises and then the analysis is carried out to support the needed decisions. Big Data analytics can be used in School

Education System of J&K for finding out solutions to certain issues like student retention, improving admissions, drop-out percentage, improving the progress of students, improving teaching and curriculum, pass-out percentage, etc. Analytics in school education can also help institutes in enhancing the students' success rate. Big Data analytics is carried out by following the steps as under:

- Firstly data is collected and managed efficiently.
- Then this data is preprocessed and cleaned.
- This data is then transformed into a proper structure.
- Analysis is done on this data using various methods and tools.
- Finally results are represented or visualized and then interpreted.

### 3.1 Types of Analytics in School Education Sector

The process of identification, measurement, and understanding the results extracting from volumes of data is known as analytics [6] and can be divided into three classes as descriptive, predictive, and prescriptive. There are two types of analytics mostly used in education sector as shown in fig.3 with extension of further types, which are described as follows:

*Learning Analytics*: It was defined in the first International Conference on Learning Analytics and Knowledge held in 2011 in Alberta as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” or it can also be defined as “the analysis and interpretation of large amounts of data generated during student learning in order to assess academic progress, predict future performance and identify potential problems” [1][4]. This type of analytics is used to improve learner's success by making decisions required for achieving specific learning outcomes and is applied at Microlevel and Nanolevel of analytics. Behavioral patterns of students can be identified by using learning analytics which intern can be used by teachers to identify risk prone students, so that appropriate preventive measures can be taken by the institution management and the administration. Learning analytics can also be used to check and prevent the dropout percentage of students, so that new policies or schemes are formulated to control or decrease the number of dropouts from the institutes because presently a high number of traditional students are decreasing in school education system of India and so is in the school education system of J&K. In 2018, almost 95% children have enrolled in primary schools, 69% in secondary and shockingly meager 25% students got enrolled in post secondary education in India [7]. Learning analytics focuses on identifying and optimizing interactions of students with the education system [8], like assessment, analysis, and presenting of learner actions (Time spent in classroom interactions, performance of student in a course, student interaction with the learning course, etc.) to understand patterns and optimize learning process.

Learning analytics models are used for processing and analyzing the data from education system, which can help School Education System of J&K to understand learning and teaching process effectively. Learning analytics helps in identifying student learning patterns and is based on student learning experience and monitoring of course, e.g. monitoring and assessing “engagement index” for students can help in

increasing student involvement in various educational activities [8]. Learning analytics through various methods identifies student performance and provides tools for continuous improvement in learning process. In J&K, institutes of school education are facing greater demands to retain students which can be fulfilled through learning analytics by providing personalized learning experience to students by using data according to needs of students, which leads to greater success in the classroom. It can be used at various levels of school education in J&K like course, curriculum, and institutional level, and a value can be extracted by using Big Data analytics at all these levels. From the point of view of teachers and managers of the institutes, learning analytics help in assessing teaching curricula, improving the school's established assessment methods, analyzing in-depth teaching data, identifying under-teaching and finally finding out better methods for teachers to help students solve practical problems [4].

Six dimensional domain of learning analytics is described as follows: [5]

- **Stakeholders:** Those interested persons who get benefits from analysis process of education data and are further grouped as the primary group who are directly interacting with teaching and learning process (e.g. teachers, students, etc.), then is secondary group who are indirectly involved with the learning process (e.g. alumni, parents, etc.) and the last but not the least is hybrid group who is concerned with the analysis and decision making processes.
- **Objectives:** These describe the main focal points or aims of an institute for carrying out analytics such as academic and administrative goals.
- **Data:** It is the central point of action which is collected and managed for further evaluation through modern Big Data tools and techniques.
- **Instruments:** These are concerned with the available tools and technologies for Big Data analysis (e.g. Hadoop, Spark, MapReduce, Oracle data mining, etc.).
- **Constraints:** Deals with those who have access to the data and also involve legal and ethical issues related to data collected like privacy, norms, timescale, etc.
- **Limitations:** These are concerned with short-comings or deficiency of institutes in having skilled and trained staff so that clients can understand visualized information efficiently and use it fully.

Presently four dimensional model for analytics is available; **Data and Environment:** relate to 'what' data is available. **Stakeholders:** deal with 'who' is targeted. **Objectives:** It is about 'why' to do the analysis. **Methods:** deal with 'how' to perform the analysis.

What → Who → Why → How  
(Data) (Target) (Analysis) (Methods)

**Academic Analytics:** It can be defined as "the combination of information, technology, culture, management for

managing the academic enterprise" [5] and involves decisions related to the campus services like financing, administration, academic counseling, etc. It is used to make decisions that help out administration and other operations of institutions, and the operations are carried out at Mesolevel and Macrolevel of analytics. Academic analytics works at an institutional level for improving effectiveness by using academic data of students and the institutional data to check the working of an institute. It analyzes non-performance and non-environmental aspects of learning process like student recruitment, student admissions, student persistence, student retention, and other administrative aspects [8]. For financial stability, institutes have to improve student admissions, retention, and their performance, which can increase learning success and engage students in the course. Academic analytics can be used in School Education of J&K to identify drop-out prone students in a course and thus promises a quick action to retain the students till completion of the course. It operates at the higher level to help administrators, funders, the national government and other big educational organizations like UNESCO.

Other types of analytics which can also be applied to school education data are as:

- **Predictive analytics:** It is concerned with the statistical analysis of data to extract useful information and is useful for behavioral and events prediction. It can be used to identify hidden relationships among data. It is being used to provide many services and products. It is used to analyze different patterns from students and courses (overall student response in the institute) like student achievement patterns, performance patterns of students, etc., rather than individual student responses. Predictive analytics can help in teaching of subject matter and provide quick feedback through some forum [8], which can significantly improve the learning among students. Predictive analytics on Big Data of School Education in J&K can improve learning standards among students and can also be used for fund raising and fund utility operations. It can also be used as a key operation in admissions improvement for school education institutes of J&K.
- **Visual analytics:** It is a multi-disciplinary domain in which analysis of large and complex datasets is carried out through automated analysis techniques with interactive visualization for accurate decision making [9]. It is carried out by combining data analysis, visualization, and human visual perception for manipulating Big Data.
- **Business analytics:** By using Big Data, education system can become a strong enterprise at both national and international levels. This type of analytics is used for effective business purposes by using data and its analyzed results which can be helpful for institutional enterprise. It is helping in improving the efficiency of many organizations.



Fig. 3: Types of Analytics for school education Big Data

According to Educational Business Construct Outlook, learning analytics, academic analytics, and educational data mining are related with each other to meet the needs of students, faculty, administrators and other stakeholders [10]. Big Data application in education is mainly for online decision-making, data mining, and learning analysis, but for School Education of J&K where online education of students is not prominent, still two of the three applications (data mining and learning analysis) can be carried out. Through these applications, study and application of forecast analysis, academic analysis, and behavior analysis is possible. Various techniques like statistics, machine learning, computational linguistics, etc. [2], can be used to carry out visualization, descriptive, and predictive analysis to extract useful information from large amounts of data to find out solutions for further enhancing the school education sector of J&K. After identifying variables from the data generated by learners in learning process, related patterns of learning can be found out by using some analytical algorithms for predicting the success or failure of students in learning. Machine learning approach can be applied to student data at individual level as well as at entire population level of students in School Education System of J&K to study student admissions, drop-out rate, and task completion by students. Big Data in School Education of J&K can be used in various ways for designing, developing and evaluating courses under the influence of human intervention. Applying data science on educational Big Data is a combination of technology and art to handle data effectively for carrying out analysis.

can be collected, managed, analyzed and the results can be reported to the students and teachers about learning progress and the learning environments. The system begins with the evaluation function which gives rise to Big Data in school education such as courses, interactive discussions, pre-test, student evaluation, teacher evaluation, etc. The data generated through all these activities is collected, managed, and stored in a storage system which can ensure the comprehensiveness of data for performing some necessary mining and analysis operations. For analysis data is transferred to analysis subsystem which is the core part of this whole system in which different kinds of analysis can be carried out on the data that can be helpful for determining academic progress, predict future performance, identify potential problems, etc. in School Education of J&K. Finally a feedback subsystem is designed to visualize the outcome results of data analysis subsystem. In the present information age, visualization is an essential part of data analysis and displaying of results. A visual tool known as “dashboard” is used here to present feedback information of the system to the learners, teachers, and the management in time so that these active participants in school education of J&K can have through understanding of the learning situation. Transmitting feedback information to teachers can enable them to give timely intervention to the learners’ learning activities when necessary. Big Data analytics can be applied at all levels, ages, and venues of school education sector in J&K for high implications [8], but it is also important to identify the opportunities and risks in this technology.

Fig.4 shows the introduction of learning analytics technique for School Education of J&K, in which student data

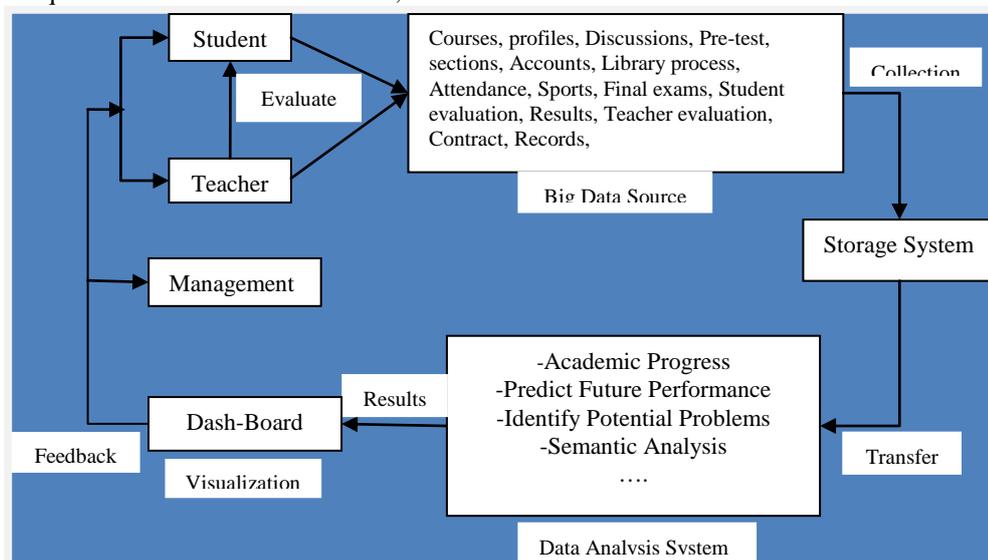


Fig.4: Flow Chart of Analysis System for School Education of J&K

#### IV. OPPORTUNITIES FOR BIG DATA ANALYTICS IN SCHOOL EDUCATION SECTOR OF J&K

There are various benefits of Big Data analytics in School Education System of J&K and also provides many opportunities of innovation in the school education in future. Students' data analysis helps in finding out the relevant issues in study process and also optimizes the study process of students, which helps them adopt good learning habits leading to the goal of self-learning. Various opportunities provided by Big Data analytics are described as follows:

- *Motivation for contribution:* Students should become aware of the benefits provided by Big Data analytics in improving performance of students, so if these students can realize the impact and benefits of Big Data themselves, then more and more interested students can input data in the process of analytics for betterment of the system. Teachers and management of institutes in School Education of J&K, where knowledge or awareness of analytics is feeble, should be provided with awareness about power of analytics in education sector also involving students in the process, so that complete data of school education sector becomes available.
- *Individual attention:* Personalization of education is possible through Big Data analytics. Educational experts can use Big Data to frame course curriculum according to the learning ability of each student enrolled in the course and also according to the cope up capacity of students. This process can be also beneficial for e-learning system in the school education system of J&K to revolutionize the learning and improve the results. Thus, through this process student engagement in the class and with the course will increase highly and hence will increase classroom success of students in school education system of J&K.
- *Efficient productivity:* Big Data can save a lot of time and effort of students in preparing for exams both competitive as well as grade exams by analyzing the data collected from students who have already appeared or passed these exams and thus helps them to prepare by understanding the patterns of study and preparation of these students for good results.
- *Cooperation:* Institutes in school education of J&K can use Big Data analytics to identify institutes at regional, national, and international levels for promoting collaboration among them and thus a competitive environment can be created for healthy education by making comparisons of work performance, teaching, methodology, etc. among these institutes.
- *Funds and Finance:* Patterns of finance or funds expenditures can be detected through Big Data analytics which will be helpful for reducing expenditures and efficient management of finance or funds at institute and administrator levels.
- *Effective Learning:* Big Data analytics makes it possible for students and teachers to make self-assessment and self-measurement of their learning improvement which can be helpful for increased success rate and knowledge flow in the School Education System of J&K. There are various

methods through Big Data analytics by which learning progress of students can be reflected and appropriate decisions can be taken about improving the learning process of students. One of the methods is personalized feedback system which is very useful but time consuming for teachers, so a timely application of this method cannot always be possible. Personal analytics can be used in School Education of J&K to identify student's common errors in understanding and using of less appropriate techniques by students to reflect their development, and to identify improvements in a timely manner.

#### V. ISSUES REGARDING BIG DATA MANAGEMENT AND ANALYSIS IN SCHOOL EDUCATION SYSTEM OF J&K

School education of J&K is facing various issues in handling Big Data, which are explained as follows:

- *Issues of Data Storage:* A huge amount of data is being generated by each and every student in different formats (structured, semi-structured, and unstructured) during the learning process or education process which is required to be stored for further management and analysis process and certain data of students needs permanent storage (lifetime of a student) for future operations like certificate verifications and other personal information checking. Due to the increasing input capacity of admissions of students in school education of J&K, storage space provided by conventional storage devices is not sufficient for the storage of data generated by the students and thus, marks a great issue for the school education system. The institutes in school education of J&K have to make important decisions about the aging factor and permanent storage of data that is, which part of data should be discarded after certain time and which part is necessary for further analysis and also about the part of data which needs permanent storage.
- *Data Management Issues:* In school education of J&K, issues of managing data of various types and formats is a major area of concern for the administration of institutes, which is mostly related to the issues of access and governance of student data. The heterogeneous data generated by students during learning process is mostly handled by different departments or sections of an institute (library, accounts, admission, etc.). For analysis this data has to be brought together and relate them which is also a challenging task in data management and after cleaning the data, some errors or incompleteness may exist which also adds some difficulties in managing this data. In J&K, School Education System has not yet adopted an electronic way of managing data at institute levels in most of the areas and still depends on manual way of managing data of students, so this data needs to be transformed into an electronic copy for further actions like submission to administration, analysis purposes, securing the data, etc.
- *Issue of Affordability:* Due to the least interest of administration and governance of School Education in J&K, automation of school records and data management at institute levels has not been initiated fully. Therefore institutes in the

school education cannot afford to manage student data or school records electronically due to the lack of funds and resources like skilled manpower, tools and technologies, etc.

- *Data processing issues:* Presently data is being generated in huge volume and variety with high frequency which cannot be handled through conventional technologies and resources in managing and processing the data. As the size of Big Data generated and collected in School Education of J&K is large, a lot of time may be consumed in processing, so appropriate tools and techniques of data management and analytics are needed for its timely processing. Modern and upto date technologies of processing are needed to detect maximum patterns out of the collected data.

## VI. CHALLENGES FOR BIG DATA MANAGEMENT AND ANALYSIS IN SCHOOL EDUCATION SYSTEM OF J&K

In School Education of J&K Big Data trend has not been started yet, but it will have to face various challenges to adopt this innovative technology like huge volume, variety of data, no proper structure, incompleteness, timeliness, privacy, security, etc. School Education System of J&K faces various challenges in using Big Data for predictability and decision making purposes. Like other sectors in J&K, school education also encounters some challenges in using Big Data due to increasing volume and variety and also some challenges arise specific to this sector only. These challenges of using Big Data in school education can be ethical and practical in nature [5], which are described as follows:

- *Personal Consent:* From the ethical point of view, student's willingness to use his/her data for analysis is important and a formal permission should be sought from the concerned students, because students do not feel comfortable sometimes in using their personal data for analysis. Some institutes outsource the task of analysis to any third party or organization, in that case student's consent should be sought and the decision of student must be respected.
- *Privacy and Security:* Institutes in School education of J&K should manage student data so carefully that the privacy and security of students should not breach, so the most personal data of students should be avoided for analysis and only academic data should be used for analysis, and if institute is using some personal data of students like health state, socio-economic status, etc., then utmost care should be taken during management and analysis about the privacy of each and every student. A student pursuing education in an institute generates a lot of data that describes his/her identity and personality and this data can be used to reveal a lot of information about him/her, which sometimes can be misused. Therefore to overcome this challenge of misusing data of students and analyzing it unauthorizingly, institutes should store and manage this data securely and thus prevent any fear of fraudulence and misuse. Examination data of students (grades and evaluations) in schools should be kept private so that students mostly stress on gaining knowledge while learning rather than getting fame or

reputation. This can help in getting qualitative education rather than quantitative education.

- *Fooling the System:* Data quality is important for getting good results from analysis, some students in the institute for the sack of fun, may manipulate the system by pretending to learn or just killing time in the library spending hours there and also issuing many books in his/her name for showoff, which can help him/her to manipulate or cheat the analysis process about his performance. Institutes are responsible for efficient data management and filtering out bogus data of students, so that precious data sets can be stored and managed securely.
- *Institutional Capability:* Institutions of School Education System of J&K should be capable to gather and manage data of students efficiently by using some technology, so that data is consistent and accurate because accuracy of analysis depends on the quality of data available. This data can be managed at institutional level or at administrative level of school education system through a strong data management system where institution can link the data available from different departments or sections of the institute whereas administration can link the data collected from various institutions of school education system. By collecting and managing data in this way can be expensive sometimes, but will be helpful for better analysis and thus for the good governance of the system.
- *Staff Capability:* As the Big Data is a new concept in School Education System of J&K, so awareness programs should be held for the employees and the management of the institutes about this new trend to make them familiar about the benefits of analytics in school education. Institutes should be aware of the advantages provided by learning analytics, so that some employees can be trained to use Big Data analytics in school education for improving teaching and learning process in institutes. Institutes and administration of school education system have to make some investments in training manpower for management and analysis of Big Data generated by students during learning process and they can also hire new skilled employees and analysts for carrying out further operations on this data.
- *Claims Beyond data:* Results of analysis can be faked by the analyst or the institute, for example for rankings of institutes at national or international levels, one institute can claim itself best on the basis of academics, while other on the basis of infrastructure and so on.

## VII. APPLICATIONS AND BENEFITS OF BIG DATA ANALYTICS IN SCHOOL EDUCATION OF J&K

Big Data Analytics have various applications in the field of School Education in J&K, which can benefit students, teachers, and the administration in the prediction and decision making by assessing the progress and performance during various educational scenarios. Big Data in School Education of J&K can also be applied for personalized learning, adaptive

learning and trend analysis, which are also the ultimate goals of Big Data application in school education. Some of the applications of Big Data analytics in school education are described as; Degree Compass [2], a kind of recommendation system inspired from recommendation systems developed by Netflix, Amazon, and Pandora, recommends students with the suitable courses after a thorough matching according to their capability by using predictive analytics on enrollment and grades data of students and ranks courses in accordance to factors that determine how helpful a particular course might be to the student to advance through the learning process. Various analytical tools like Tableau, Quibble, Qlike, etc. can be used for analysis of school education data. Another technology known as Electro Encephalography (EEG) sensor can be used to measure brain electric activity of students in School Education of J&K for checking attention or focus level of students of a particular group or section during lectures, as this technology is available at affordable prices. Predictive analysis can be carried out on the data generated through these sensors using Big Data technologies. The learning service platforms can be developed for School Education in J&K, through which data can be collected about individual behaviors of students in the form of search paths, different learning contents, recording and feedback data, which can be analyzed latter on.

For School Education in J&K, Big Data possesses great opportunities which remain unattempted and unexplored due to the lack of implementation of Big Data management and analytical processes and policies. In School Education of J&K, Big Data analytics can be applied to a wide range of operations at institution as well as administration level such as dealing with enrollment of students, marketing, finance, teaching faculty, performance checking, human resources and facilities provided by the institute, to gain successful results for improvement in imparting education. In addition to J&K, improving admissions, performance and preventing student drop-outs will be important in school education system at national level too. Here are some of the promising benefits of Big Data management and analytics in School Education of J&K:

- **Admission Process:** Communication tools that are being used to reach out to the interested candidates for seeking admission in a course have changed from traditional media like newspapers, magazines, newsletters, etc. to online advertising and social media, which is being used by various educational institutes for marketing purposes, for example Harvard [2], the university of California and other institutes throughout the world are using social media as an advertizing tool. Presently students before getting enrolled for any course or making any contact with the institute carry out extensive research and discussion through online means like forums to check the individual sentiments of students about any particular institute or course for reaching on a final conclusion of admission. Factors like academic quality, facilities, campus, financial aid, etc. play an

important role in student interest for choosing an institute or course. Student enrollment can be enhanced for an institute by analyzing the sentiments of students through one of the Big Data solutions called sentimental analysis, thus can serve as insight analysis platform for school education. Big Data analytics is used by various organizations to assign ranks to the institutes on certain criteria at national and international level which can help students to take quick decisions in choosing an efficient institute for admission.

Various school educational institutes, mostly private schools, in J&K are using social media as an advertising tool to promote and fetch maximum admissions from the territory. In order to acquire maximum number of admissions in a session, Big Data analytics can be applied on applications received, test scores and interviews held, which can also help in filtering process of screening students. Students can make many options of applying for multiple institutes at a time which is a tough job for an institute to forecast the choice of a student that can be solved through Big Data analytics for fetching maximum number of admissions. Demographic as well as historical data [5] of students can be used by institutes for analysis to make decisions regarding the admissions in a particular area. By knowing the demands of students using Big Data analytics, institutes can make improvements in the existing curricula, teaching methodologies, and other facilities provided, thus attracting more and more efficient students for admission in the institute.

- **Student Performance:** It is largely assessed through the students' progress in academics. Student's conduct can be measured through analysis and assessment of student's overall campus life like tasks provided, social habits, eating and sleeping habits, effectiveness of the tutor and so on [2]. IOT technologies such as sensors can be used inside school premises to track the academic life of students by checking their schedule of time spending in classrooms, library, hostels, cafeteria and so on. Thus, the large data sets generated through these IOT technologies can be analyzed to get the desired results. For forecasting student's performance in school education of J&K, factors like student profile, history of performances, and demographic information are helpful and should be kept in mind during analysis process. Big Data technologies have made it possible to collect, store, manage, and process the large amounts of data with wide variety and fast generation time and thus these tools can be used extensively to analyze and assess the different characteristics of students involved in the learning process of school education in J&K. Students can be provided regular and immediate feedback regarding their progress and performances through Big Data analytics, which is an efficient and reliable method for teachers to check the overall performances of students. Big Data based feedback system is better than the traditional methods for checking the progress of students in school education system and also contributes in the development of an efficient competitive environment among the students for the growth and progress of their learning.
- **Preventing Student Drop-outs:** One of the biggest problems in School Education of J&K is dropout percentage of students from schools,

which is increasing day by day due to various reasons, thus affecting the education environment in J&K. Table-VI shows the drop-out percentage of students upto secondary level in J&K during the year 2009-10 [3]. Loosing students from institutes in this way draws negative effects on the institute strategies and financial planning, thus institutes are not able to maintain the quality education standards. The students showing possibility of dropping out from the institute can be identified and necessary measures can be taken by using the data of interaction of those students with the course and the institute, which can show their engagement and involvement with the institute such as falling grades of a student, low attendance, less interaction with his fellow students, etc. [5]. Thus, Big Data analytics can be used to identify the students at risk in order to decrease the student drop-out percentage. Psychologically upset or depressed students can also be identified through Big Data analytics, which can help in preventing student suicides and thus behavioral interpretation of students can help the institutes or administration to monitor and control students in a better way.

**Table VI: Dropout percentage of students in J&K, 2009-10**

State	Class	Boys (%)	Girls (%)	Average (%)
J&K	1-5th	9.8	6.8	8.4
	1-8th	23.8	8.19	22.0
	1-10th	47.2	40.6	44.3

- **Enhancing Teaching Efficiency:** Big Data driven education has changed the role of teachers from being a mentor to a technical educational analyst. The main motive behind the usage of Big Data in School Education of J&K is to formulate better teaching strategies through data mining and learning analysis which can be used for optimization of teaching strategies and also can improve student's learning efficiency. Teachers can plan methodologies of teaching a course by identifying the student requirements, learning curve, and future performance [2] through timely feedback system which is possible by using Big Data technologies. Institutes can develop feedback system for teachers by using various factors of teaching methodologies like teaching through projectors or whiteboards, applied or practical application knowledge, analyzing recorded lectures, etc. Student sentiments can help teachers and the organization to identify the standard of study material and views about the teachers concerned, which intern can help them to take corrective measures or actions. Through this process successful teachers can be identified and rewarded who have gained much positive sentiments from students. For selecting a teacher for a particular class or section according to the requirements of students of that class is possible through analyzing the data like number of students, students' behavior, students' ability, etc.
- **Student Oriented Courses and Employment Avenues:** Students in School Education of J&K, after qualifying high school (10th) can choose their major subject or stream for secondary education (10+2) through Big

Data analytics, which can help them in understanding the future avenues of a particular stream or subject and also utilize their investment and ability to the fullest. Big Data analytics can be used here to get the appropriate information of institutes and make due decision of selection of course in an institute. This analysis can be carried out on the collected data from various sources like institutes, alumni feedbacks, and from students and parents about their expectations from the educational institutes. Big Data analytics can help students to identify their hold on the course and the required skill-sets to qualify any competitive exams for choosing a profession; on the other hand companies or organizations can also use Big Data for identification of suitable candidates for their working process. So by using these methods success rate of students in competitive exams can increase and students can get admission in desired courses, in return the colleges or training institutes can get best and most deserving candidates for excelling in their professional life.

- **Effective administration system:** Identifying needs of students at administration and management level of school education through Big Data analytics can help administration to do needful for the students and school welfare. To compete in a globally competitive environment of education system, institutes or administration should have a complete information about marketing trends and behavioral patterns of learners, which requires analytical and intelligence architecture [11].

## VIII. BIG DATA MANAGEMENT AND STORAGE IN SCHOOL EDUCATION SYSTEM OF J&K

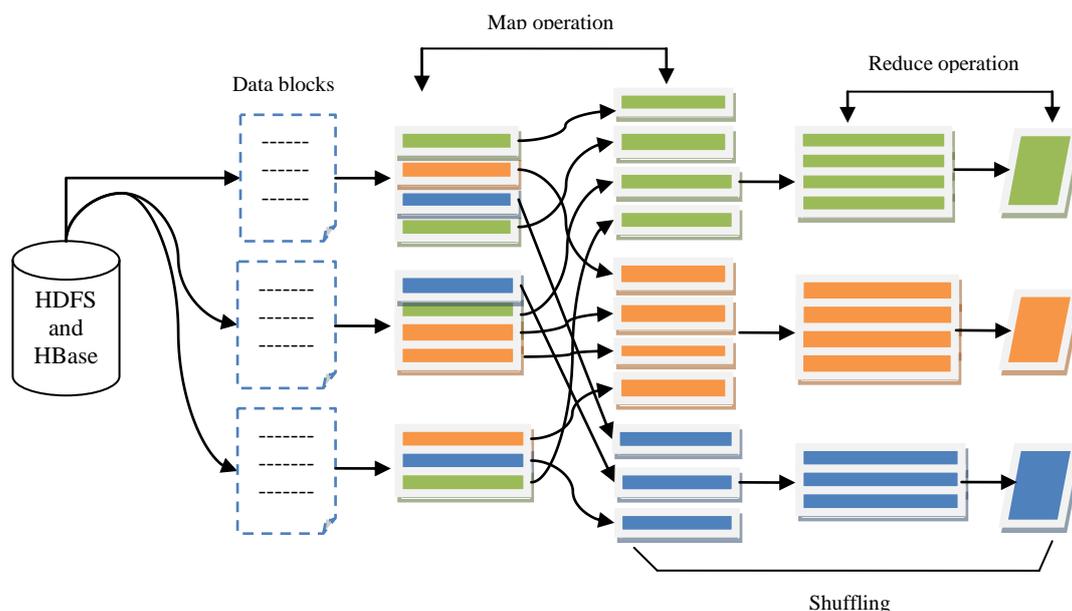
As it is evident from the above discussions that Big Data management and analytical concepts can be applied to School Education of J&K at administrative as well as institutional levels which includes admissions, recruitment, student performance tracking, financial planning, etc. Structured data through traditional technologies is collected, normalized, and stored in databases, which is then used as a main resource for manipulation to support an organization or a system, but the unstructured or semi-structured data generated alongside is generally in massive quantity in comparison to the structured data, which cannot be efficiently processed through conventional technologies, thus left aside or trashed unattended. Also there are various challenges arising in Big Data analysis due to the huge volume, velocity, and variety of Big Data. Thus some distributed computing frameworks were developed to cater this demand [12], e.g. MapReduce introduced by Google (Batch Processing) and Dryad introduced by Microsoft (Data flow process). Nowadays Big Data management and analytics is usually associated with various modern technologies like Hadoop, NoSQL, Spark, Cloud Computing, etc., because of large data sets and huge volume of data to be processed in time. With the advent of these new technologies, this complex data which is in a big quantity called as Big Data has now become a most valuable resource for any organization or system. The storage and processing costs of these Big Data technologies is much lower as compared to conventional technologies in storing and

processing of large amounts of data, thus is an efficient platform for data storage and analysis.

The School Education of J&K is lacking necessary facilities for data management and analysis such as computational capabilities, tools, and human resources, and thus institutes or the administrators are not able to utilize the Big Data prominently as compared to other domains or systems in J&K. Protecting privacy and confidentiality of students and providing high level of security to the student data can be one of the reason which makes hindrance to pace up the utilization of Big Data technologies in School Education of J&K. For prediction of student sentiments and harnessing of various aspects of School Education of J&K through modern technologies of data mining is essential. Due to the generation of large amounts of data through student learning process and increasing amount of semi-structured and unstructured data generation in School Education of J&K large amounts of storage space is required for Big Data. So a distributed data processing model based infrastructure is needed for storage, processing and analysis of Big Data arising from School Education of J&K, which requires configuration of several nodes with each other in a Master-Slave architecture in which each node has its own storage and processing capabilities and are connected to a master node which is responsible for managing this whole cluster of nodes. The interconnection of these nodes in a cluster allows massive sharing and searching of data. Hadoop is one such platform for Big Data storage and processing in a distributed fashion which can be adopted by School Education of J&K for its cost effective and efficient data management and analysis.

Various commercial areas have gained benefits by using the Hadoop technology. Hadoop being an open source platform provides solutions to various problems arising through Big Data and is also used by some of the pioneers like Amazon, Facebook, Yahoo, etc. Due to its attractive features for data storage on a cluster like reliability, scalability, and

fault-tolerance, it is being used in various fields of research [13], such as social networking, healthcare, bioinformatics, etc. Hadoop consists of various components of which two main components are: Hadoop Distributed File System (HDFS) and MapReduce, in which HDFS is a fault tolerant storage system for distributed storage while as MapReduce is performing parallel computing. Fig.5 shows complete operation of MapReduce task in a Hadoop cluster. HDFS is an open source implementation of GFS (Google File System) and MapReduce which produces high performance with commodity hardware. A large amount of data is stored in HDFS which can scale up incrementally and preserves data from losing by surviving the failures. Next component of Hadoop ecosystem is HBase, which is a NoSQL technology and acts as a database on Hadoop platform in a distributed data processing environment [14]. NoSQL technology is based on shared nothing architecture [1], in which clusters are not dependent on each other for update operations while sharding data repository that is, if any change occurs inside one cluster other clusters should not get affected. These databases provide scalability and improved performance as compared to conventional databases and are efficient for storing unstructured data, examples HBase, MongoDB, Cassandra, etc. Hive is another component of Hadoop ecosystem which has the capability to process Peta-Bytes of data [12] and acts as a data warehouse on top of Hadoop platform. HiveQL or HQL is its query language that can also perform SQL operations, thus provides an interface to the users familiar with relational data stores. Sqoop is one of the tools used in Hadoop environment for importing and exporting data from databases to HDFS and vice-versa. Hadoop functions by setting up clusters of machines, which can be low-cost computers, by coordinating tasks among them and if one machine fails in a cluster, Hadoop continues its operations without any interruption or data loss by transferring operations to remaining machines in the cluster and also minimizes the transportation cost of data [15].



**Fig. 5: MapReduce operation in Hadoop cluster**

A hybrid data storage and management system can be used for Big Data management in School Education System of J&K, which will be an integration of relational (RDBMS) and non-relational (NoSQL) databases. Different educational institutes in School Education System of J&K cannot afford Hadoop cluster management and maintenance at individual levels, so they can use various cloud computing models for Big Data storage, because storing data in cloud can be a viable option for these institutes to carry out Big Data analysis process. Presently various small to medium business organizations have grown their trust in many cloud computing models which are maturing and improving in security and data integration. Many IT companies are also supporting to deliver cloud services, thus becoming a reality for these types of organizations to use cloud computing in which private cloud deployment is more than any other cloud type, for example Amazon cloud services can be used in School Education of J&K for data analytics, which provides both hardware and software services with step by step guidance through setup formation, analysis, and result presentation.

Fig.6 depicts a clear picture of Hadoop ecosystem handling Big Data (Structured, Semi-structured, and Unstructured data) for efficient storage, management, processing, and analysis. As it is evident that for Big Data analytics and data mining, data can be integrated from Hadoop and other data stores like RDBMS, NoSQL databases, etc. through Big Data connectors for any analysis purposes and knowledge of SQL and statistical skills can be reused to apply data mining techniques for prediction of future learning behavior of students and their results. Finally developers can write Hadoop applications to access the data from Hadoop environment and also to visualize the results of analysis. Results of the operations calculated and predicted through analysis module can be visualized through any visualization tool to all types of users

(teachers, students, management, and administration). Fig.7 gives the detailed interactions among different parts of this proposed Big Data management and analytical system for School Education System of J&K. The process begins with generation of various kinds of data by students through their interaction with the course opted. This data (course data, student data) is stored and managed through Hadoop architecture with HBase (Distributed data storage) or any other NoSQL database. The analysis or prediction is done on the data collected from the Hadoop clusters, NoSQL database and Information Management System using various analytical tools and models available. The results obtained from analysis and data mining are presented in visualized formats like graphs, plots, etc. (Bar graphs for student's progress, Histograms representing students' problems with particular activity, Tree-structure for knowledge, etc.) which can be accessible to the teachers, principals, management, administration, and also to students. These results can provide students with suitable learning guidance and learning strategies and can also guide teachers and management to take appropriate steps in enhancing student performance and teaching methodology through their commands and instructions. The end results obtained through analysis and forecasting done on the students' data can be used to review the syllabus of curriculum and content of subjects according to the students' learning level and their interest in the course. Finally Human intervention of teachers, principals, and managers is needed on the basis of forecasted results to check various kinds of behaviors of students within the campus as well as in the society. School education in J&K is mostly delivered under the constraints and control of teachers, so that a constant supervision can be held on the students' learning process and thus students can be lost in learning. This process of timely and constant supervision in the learning process of students collects and analyzes the student data, displaying the deficiencies or major problems in the learners' and learning system by providing the feedback and also suggesting the remedy advices to improve.

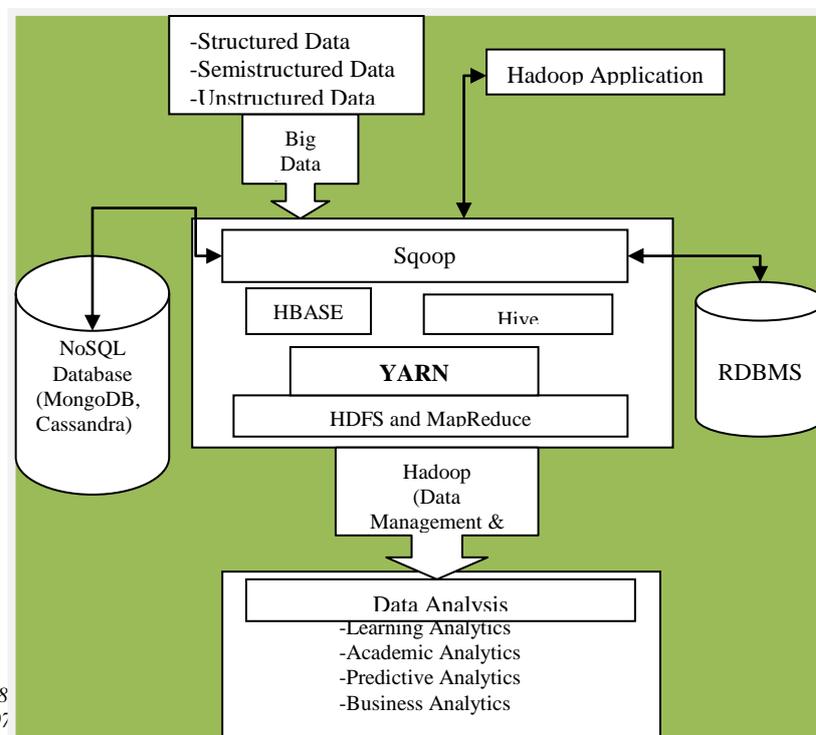


Fig.6: Framework for Handling School Educational Big Data of J&K through Hadoop

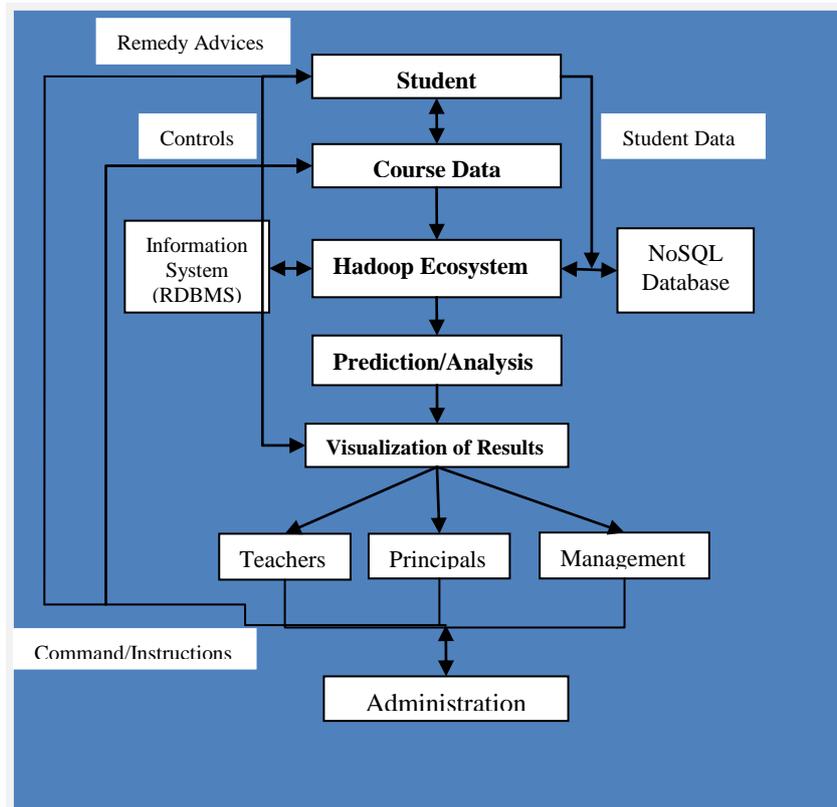


Fig.7: Flow Chart of Big Data management and analysis system for School Education of J&K

## IX. CONCLUSION AND FUTURE WORK

We identified the sources of Big Data and its power to improve the condition of School Education in J&K. The value of Big Data analytics can be harnessed to revolutionize the School Education System in J&K and thus set up new goals and methods in learning process. New teaching and learning methods can be introduced in school education by using Big Data analytics, thus improving the quality of teaching, learning and the services provided by institutes and can also help in formulating effective management model for school education. Big Data generated in various school educational institutes can be gathered, stored and managed in a centralized way, so that these institutes will get unified under a single administration which can carry out analysis on this data based on common goals of these institutes. To find out solutions for issues and challenges of Big Data some smart policies should be drafted by the administration and governance of School Education System of J&K. Ethical

challenges of using student data in School Education of J&K like privacy and students' willingness should be handled by institutes or the administration through secured storage and processing of adequate and relevant data gathered for specific purposes. Sensitive data of students and institution can be managed separately through better data management systems which can provide high security and protection of data through encryption means. The issues of learning, admissions, student retention, and recruitment of qualified and trained faculty in the School Education System of J&K can be solved and improved with the help of Big Data technologies. IOT devices like sensors can be installed in the campus of the institute to collect behavioral data of students easily. The administration or governing body of School Education System in J&K can collect and manage the data from various coordinating or constitutional institutes and then perform different analytical operations on this data in coordination with any IT company which will implement an efficient data management practices and deploy better analytical tools on the Big Data analysis process for these institutes.

In future, we will carry out an implementation of this proposed system of Big Data management and analysis for School Education of J&K where we will use a hadoop multimode cluster with HBase for data

management and MapReduce for parallel computing and analysis. The data will be used from School Education System of J&K.

## REFERENCES

- [1] Ashraf, H. M. El-Bakry et al, "Enhancing Big Data Processing in Educational Systems," *Advances in Computers and Technology for Education*, 2014, pp.176-181.
- [2] Rao and K. Baglodi, "Role of Big Data in Education Sector: A Review," *International Journal of Advances in Science Engineering and Technology*, vol. 6, issue 1,spl. Issue 1, Feb. 2018.
- [3] F. Fayaz and S. Mehta, "Analysis of Education Sector-Study of Kerala and Jammu & Kashmir," *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, vol. 23, issue 3, March 2018, pp. 44-51.
- [4] Y. Li and X. Zhai, "Review and Prospect of Modern Education Using Big Data," *Procedia Computer Science*, ELSEVIER, vol. 129, 2018, pp. 314-347.
- [5] Vatsala, R. Jadhav and Sathyaraj, "A Review of Big Data Analytics in Sector of Higher Education," *International Journal of Engineering Research and Application*, vol. 7, issue 6, part 2, June 2017, pp. 25-32.
- [6] A. Manohar, P. Gupta et al, "Utilizing Big Data Analytics to Improve Education," *Semantic Scholar*, 2016. Available: <https://www.semanticscholar.org/paper/Utilizing-Big-Data-Analytics-to-Improve-Education-Manohar-Gupta/1792463683531cd5db886972d96b92aea75db18e>
- [7] P. Dar, "Top 7 Sectors where Data Science can transform India," August 16, 2018. Available: <https://www.analyticsvidhya.com/blog/2018/08/top-7-sectors-where-data-science-can-transform-india-with-free-datasets/>
- [8] A. Peterson, "Big Data in Education: New Efficiencies for Recruitment, Learning, and Retention of Students and Donors," *Handbook of statistical Analysis and Data Mining Applications*, Academic Press ELSEVIER, United Kingdom, 2<sup>nd</sup> ed., chap. 13, 2018, pp. 259-278.
- [9] J. Kohlhammer, D. Keim et al, "Solving Problems With Visual Analytics," *Procedia Computer Science*, ELSEVIER, vol.7, 2011, pp. 117-120.
- [10] L. Khanna, S. N. Singh, and M. Alam, "Educational Data mining and its Role in Determining Factors Affecting Students Academic Performance: A Systematic Review," In *Proceedings of 1<sup>st</sup> India International Conference on Information Processing (IICIP)*, IEEE, Aug. 2016, pp. 1-7.
- [11] U. Kumar, "Bringing Big Data to Your School's Analytics," *Higher Education Marketing*, September 3, 2013. Available: <http://www.highereducationmarketing.com/blog/bringingbigdataschoolsanalytics>
- [12] Q. Liu, Y. Fu, G. Ni and J. Mei, "Big Data Management Performance Evaluation in Hadoop Ecosystem," In *Proceedings of 3rd International Conference on Big Data Computing and Communications (BIGCOM)*, IEEE, 2017, pp. 413-421.
- [13] W. T. Wu, W. W. Lin et al, "Energy-efficient hadoop for big data analytics and computing: A systematic review and research insights," *Future Generation Computer Systems*, ELSEVIER, vol. 86, sep. 2018, pp. 1351-1367.
- [14] C.T. Yang et al, "Implementation of Data Transform Method into NoSQL Database for Healthcare," In *proceedings of International Conference on parallel and Distributed Computing, Applications and Technologies*, IEEE, Dec. 2013, pp. 198-205.
- [15] N. V. Patil and T. Patel, "Apache Hadoop: Resourceful Big Data management," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, special issue 4, April 2014, pp. 201-209.

## AUTHORS PROFILE



**Jahangir Kamal** is a Ph.D. Scholar in Information Technology at Jagannath University Jaipur, Rajasthan. He received his master's degree in Information Technology (MSc. IT.) from BGSBU Rajouri, J&K in 2011. He has two paper publications in national and international conferences, and one in an international Journal. His research interests are Big Data, Distributed Computing and Open-Source Technologies.



**Prof. (Dr.) Meenu Dave**, M.Tech., Ph.D. (Computer Science) has taught Computer Science in different capacities at a number of Engineering Colleges and Institutes. She has extensive experience in teaching Cloud Computing, Artificial Intelligence, Knowledge Management and Data Mining at the post graduate level. She has also authored several research papers in the specified areas which have been published in leading journals.