GUJARAT TECHNOLOGICAL UNIVERSITY

SUBJECT NAME: Resource Management for scalable system
SUBJECT CODE: 715505
Semester I

Type of course: NA

Prerequisite:

1. Compiler Design
2. Bigadata Analytics
3. Data mining

Rationale: NA

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks				Total Marks
				Theory Marks		Practical Marks		
L	T	P	C	ESE(E)	PA (M)	ESE	PA(I)	
3	2	0	4	70	30	30	20	150

L- Lectures; T- Tutorial/Teacher Guided Student Activity; P- Practical; C- Credit; ESE- End Semester Examination; PA- Progressive Assessment;

Content:

Sr. No.	Content	Total Hrs	% Weightage
1	Fundamental need for scalable systems, Limitations in hardware and software using for building scalable systems, Introduction of Google File System, Google's MapReduce, Hadoop, Amazon's Dynamo, ZooKeeper, Storm.	10	20
2	Interactive data exploration, Scalable analysis and introductory machine learning on Hadoop Applications, introduction to sequence analysis, Analytics workloads in large web companies, Graph data processing, Peta-Scale Graph Mining, Distributed transactional storage, advanced sequence analysis, Introduction to scalable convex optimization, Parallel and randomized matrix algorithms.	10	25
3	Apache Mahout, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, MapReduce: a flexible data processing tool, Map-Reduce for Machine Learning on Multicore, MapReduce and parallel DBMS	10	25

Reference Books:

1. Professional Hadoop Solutions Boris Lublinsky, Kevin T. Smith, Alexey Yakubovich ISBN: 978-1-118-61193-**7**

Course Outcome:

After learning the course the students should be able to:

1. Demonstrate the basic structure of scalable system
2. Discuss in detail about the real time file system and its analysis
3. Study in details about the large scale data analysis
4. Study the software ‘s which is used for scalable machine learning
5. Study the different ways of map reduce functions