



- Constituent College of JSS Science and Technology University
- Approved by A.I.C.T.E
- Governed by the Grant-in-Aid Rules of Government of Karnataka
- Identified as lead institution for World Bank Assistance under TEQIP Scheme

**LESSON PLAN****Faculty: Dr. Anil Kumar K M****Subject: Big Data Analytics**

Sl. No	Topics covered	Course Outcome	Remarks
1	Introduction To Big Data: What Is Big Data? Is The "Big" Part Or The "Data" Part More Important?	CO1	
2	How Is Big Data Different? How Is Big Data More Of The Same? Risks Of Big Data -Why You Need To Tame Big Data	CO1	
3	The Structure Of Big Data- Exploring Big Data, Most Big Data Doesn't Matter- Filtering Big Data Effectively	CO1	
4	Mixing Big Data With Traditional Data- The Need For Standards- Today's Big Data Is Not Tomorrow's Big Data.	CO1	
5	Web Data: The Original Big Data - Web Data Overview	CO1	
6	Web Data: The Original Big Data - Web Data Overview		
7	What Web Data Reveals -Web Data In Action?	CO1	
8	What Web Data Reveals -Web Data In Action?	CO1	
9	A Cross-Section Of Big Data Sources And The Value They Hold.	CO1	
10	A Cross-Section Of Big Data Sources And The Value They Hold.	CO1	

11	A Cross-Section Of Big Data Sources And The Value They Hold.	CO1	
12	Case Studies: Importance of big data and its applications	CO1	
13	Evolution Of Analytic Scalability – Convergence – Parallel Processing Systems	CO2	
14	Cloud Computing – Grid Computing – Map Reduce	CO2	
15	Enterprise Analytic Sand Box – Analytic Data Sets	CO2	
16	Enterprise Analytic Sand Box – Analytic Data Sets – Analytic Methods	CO2	
17	Analytic Tools – Cognos – Microstrategy - Pentaho	CO2	
18	Analytic Tools – Cognos – Microstrategy - Pentaho	CO2	
19	Analysis Approaches – Statistical Significance – Business Approaches	CO2	
20	Analytic Innovation – Traditional Approaches – Iterative	CO2	
21	Introduction To Streams Concepts, Stream Data Model And Architecture	CO3	
22	Introduction To Streams Concepts, Stream Data Model And Architecture	CO3	
23	Stream Computing, Sampling Data In A Stream	CO3	
24	Filtering Streams, Counting Distinct Elements In A Stream	CO3	
25	Filtering Streams, Counting	CO3	

	Distinct Elements In A Stream		
26	Counting Distinct Elements In A Stream, Estimating Moments	CO3	
27	Counting Distinct Elements In A Stream, Estimating Moments	CO3	
28	Counting Oneness In A Window, Decaying Window	CO3	
29	Realtime Analytics Platform(RTAP) Applications	CO3	
30	Case Studies, Real Time Sentiment Analysis, Stock Market Predictions.	CO3	
31	Mining Frequent Itemsets - Market Based Model	CO4	
32	Mining Frequent Itemsets - Market Based Model	CO4	
33	Apriori Algorithm – Handling Large Data Sets In Main Memory	CO4	
34	Apriori Algorithm – Handling Large Data Sets In Main Memory	CO4	
35	Limited Pass Algorithm – Counting Frequent Itemsets In A Stream	CO4	
36	Limited Pass Algorithm – Counting Frequent Itemsets In A Stream	CO4	
37	Clustering Techniques – Hierarchical – K- Means – Clustering High Dimensional Data	CO4	
38	Clustering Techniques – Hierarchical – K- Means – Clustering High Dimensional Data	CO4	
39	CLIQUE And PROCLUS – Frequent Pattern Based Clustering Methods	CO4	
40	CLIQUE And PROCLUS – Frequent Pattern Based Clustering	CO4	

	Methods		
41	Clustering In Non-Euclidean Space – Clustering For Streams And Parallelism.	CO4	
42	Clustering In Non-Euclidean Space – Clustering For Streams And Parallelism.	CO4	
43	Mapreduce – Hadoop framework	CO5	
44	Mapreduce – Hadoop framework	CO5	
45	Hive, Mapr – Sharding – Nosql Databases	CO5	
46	Hive, Mapr – Sharding – Nosql Databases	CO5	
47	S3 - Hadoop Distributed File Systems	CO5	
48	S3 - Hadoop Distributed File Systems	CO5	
49	Visualizations - Visual Data Analysis Techniques	CO5	
50	Interaction Techniques; Systems And Applications	CO5	
51	Interaction Techniques; Systems And Applications	CO5	
52	Case Studies	CO5	
53	Case Studies	CO5	

Text Books:

1. Bill Franks, Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with advanced analytics, John Wiley & sons, 2012.
2. Anand Rajaraman and Jeffrey David Ullman, Mining of Massive Datasets, Cambridge University Press, 2014.

Reference Books:

1. Paul Zikopoulos, Chris Eaton, Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw Hill Professional, 2011.

2. Michael Berthold, David J. Hand, Intelligent Data Analysis, Springer, 2007.
3. Glenn J. Myatt, Making Sense of Data, John Wiley & Sons, Pete Warden, Big Data Glossary, O'Reilly.
4. Alex Holmes "Hadoop in Practice", Manning Press, Dreamtech Press.
5. Dan McCreary and Ann Kelly "Making Sense of NoSQL" – A guide for managers and the rest of us, Manning Press
6. Chuck Lam, "Hadoop in Action", Dreamtech Press

Plan of action

- **Continuous Internal Evaluation process will be conducted for 50 marks**

3 tests and 2 events will be conducted

TEST 1	EVENT 1	TEST 2	EVENT 2	TEST 3	Total
20 marks	Quiz (20 marks)	20 marks	Quiz (20 marks)	20 marks	50 marks

Teaching Methodology: Black board, Multimedia projector, Digital smart board

Signature of staff

Signature of H.O.D