

Ref No:

SRI KRISHNA INSTITUTE OF TECHNOLOGY, BANGALORE



COURSE PLAN

Academic Year 2018-19

Program:	B E – Computer Science and Engineering
Semester :	8
Course Code:	15CS82
Course Title:	Big Data Analytics
Credit / L-T-P:	4 / 4-0-0
Total Contact Hours:	50
Course Plan Author:	DANANJAY

Academic Evaluation and Monitoring Cell

< SKIT ,>
 < Bangalore – 560090, ,Karnataka INDIA>
 <Phone / Fax :+91-STD- 080-23721477 >
 <Web: dhcs.skitblr@gmail.com >

Table of Contents

A. COURSE INFORMATION.....	4
1. Course Overview.....	4
2. Course Content.....	4
3. Course Material.....	5
4. Course Prerequisites.....	6
5. Content for Placement, Profession, HE and GATE.....	6
B. OBE PARAMETERS.....	6
1. Course Outcomes.....	6
2. Course Applications.....	7
3. Mapping And Justification.....	7
4. Articulation Matrix.....	8
5. Curricular Gap and Content.....	8

6. Content Beyond Syllabus.....	9
C. COURSE ASSESSMENT.....	9
1. Course Coverage.....	9
2. Continuous Internal Assessment (CIA).....	10
D1. TEACHING PLAN - 1.....	10
Module - 1.....	10
Module - 2.....	11
E1. CIA EXAM – 1.....	12
a. Model Question Paper - 1.....	12
b. Assignment -1.....	13
D2. TEACHING PLAN - 2.....	16
Module - 3.....	16
Module - 4.....	17
E2. CIA EXAM – 2.....	18
a. Model Question Paper - 2.....	18
b. Assignment – 2.....	19
D3. TEACHING PLAN - 3.....	22
Module - 5.....	22
E3. CIA EXAM – 3.....	24
a. Model Question Paper - 3.....	24
b. Assignment – 3.....	24
F. EXAM PREPARATION.....	26
1. University Model Question Paper.....	26
2. SEE Important Questions.....	28
G. Content to Course Outcomes.....	29
1. TLPA Parameters.....	29
2. Concepts and Outcomes:.....	30

Note : Remove “Table of Content” before including in CP Book

Each Course Plan shall be printed and made into a book with cover page

Blooms Level in all sections match with A.2, only if you plan to teach / learn at higher levels

A. COURSE INFORMATION

1. Course Overview

Degree:	BE	Program:	CS
Semester:	8	Academic Year:	2018- 2019
Course Title:	Big Data Analytics	Course Code:	15CS82
Big Data Analytics			

Credit / L-T-P:	4 / 4-0-0	SEE Duration:	180 Minutes
Total Contact Hours:	50 Hours	SEE Marks:	80 Marks
CIA Marks:	20 Marks	Assignment	1 / Module
Course Plan Author:	Dhananjaya V	Sign ..	Dt:
Checked By:		Sign ..	Dt:
CO Targets	CIA Target :	SEE Target:

Note: Define CIA and SEE % targets based on previous performance.

2. Course Content

Content / Syllabus of the course as prescribed by University or designed by institute. Identify 2 concepts per module as in G.

Module	Content	Teaching Hours	Identified Module Concepts	Blooms Learning Levels
1	Hadoop Distributed File System Basics, Running Example Programs and Benchmarks, Hadoop MapReduce Framework, Map Reduce Programming	10 (5.5)	-Storage, -processing Compression's,	UnderstandL2, Apply L3
2	Essential Hadoop Tools, Hadoop YARN Applications, Managing Hadoop with Apache Ambari, Basic Hadoop Administration Procedures	10 (5.5)	-Essential s Tools, -Procedures	Apply L3,
3	Business Intelligence Concepts and Application, Data Warehousing, Data Mining, Data Visualization	10 (5.5)	-Data-mining, -corporate data	Apply L3, Apply L3
4	Decision Trees, Regression, Artificial Neural Networks, Cluster Analysis, Association Rule Mining	10 (5.5)	-Regression, -Trees corporate	Apply L3,
5	Text Mining, Naïve-Bayes Analysis, Support Vector Machines, Web Mining, Social Network Analysis	10 (5.5)	Text, Web script Social network	Apply L3,
-	Total	54	-	

3. Course Material

Books & other material as recommended by university (A, B) and additional resources used by course teacher (C).

1. Understanding: Concept simulation / video ; one per concept ; to understand the concepts ; 15 – 30 minutes

2. Design: Simulation and design tools used – software tools used ; Free / open source

3. Research: Recent developments on the concepts – publications in journals; conferences etc.

Mod ules	Details	Chapt ers in book	Availabilit y
A	Text books (Title, Authors, Edition, Publisher, Year.)	-	-
1.2	Douglas Eadline, "Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem", 1st Edition, Pearson Education, 2016. ISBN-13: 978-9332570351	3. 4.5.6.7.8	In Dept
3.4.5	Anil Maheshwari, "Data Analytics", 1st Edition, McGraw Hill Education, 2017. ISBN-13: 978-9352604180	2-12	-
B	Reference books (Title, Authors, Edition, Publisher, Year.)	-	-
	Tom White, "Hadoop: The Definitive Guide", 4th Edition, O'Reilly Media, 2015. ISBN-13: 978-9352130672	9.10	In Lib
	Boris Lublinsky, Kevin T. Smith, Alexey Yakubovich, "Professional Hadoop Solutions", 1st Edition, Wrox Press, 2014. ISBN-13: 978-8126551071	4.5	Not Available
	Eric Sammer, "Hadoop Operations: A Guide for Developers and Administrators", 1st Edition, O'Reilly Media, 2012. ISBN-13: 978-9350239261	14.16	In lib
C	Concept Videos or Simulation for Understanding	-	-
C1	Hadoop Distributed file system https://www.youtube.com/watch?v=GJYEsEEfjvk :13minutes		
C2	Storage and Processing https://www.coursera.org/lecture/big-data-introduction/mapreduce-simple-programming-for-big-results-pL4NH		
C3	Map reduce programming https://www.coursera.org/lecture/big-data-introduction/mapreduce-simple-programming-for-big-results-pL4NH		

C4	Bussiness intelligence https://www.coursera.org/courses?query=text%20business%20intelligence		
C5	Map reduce programming https://www.coursera.org/lecture/big-data-introduction/mapreduce-simple-programming-for-big-results-pL4NH		
C6	Map reduce programming https://www.coursera.org/lecture/big-data-introduction/mapreduce-simple-programming-for-big-results-pL4NH		
C7	Map reduce programming https://www.coursera.org/lecture/big-data-introduction/mapreduce-simple-programming-for-big-results-pL4NH		
C8	Data warehousing https://www.coursera.org/courses?query=text%20warehousing		
C9	Data mining https://www.coursera.org/courses?query=text%20mining		
C10	Text mining https://www.coursera.org/courses?query=text%20mining		
	Lab :		
D	Software Tools for Design	-	-
1	Not required		
E	Recent Developments for Research	-	-
1	TRP Calculation		
2	mutual funds		
3	Data warehousing		
F	Others (Web, Video, Simulation, Notes etc.)	-	-
1	Big data annalytics fdp program data enclosed		
?			

4. Course Prerequisites

Refer to GL01. If prerequisites are not taught earlier, GAP in curriculum needs to be addressed. Include in Remarks and implement in B.5.

Students must have learnt the following Courses / Topics with described Content . . .

Modules	Course Code	Course Name	Topic / Description	Sem	Remarks	Blooms Level
1	15CS82	Artificial Intelligence	Artificial Neural networks	8		L2
2	15CS82	Data warehousing	Corporate data	8		L2
3	15CS82	DBMS	quaries	8		L2

5. Content for Placement, Profession, HE and GATE

The content is not included in this course, but required to meet industry & profession requirements and help students for Placement, GATE, Higher Education, Entrepreneurship, etc. Identifying Area / Content requires experts consultation in the area.

Topics included are like, a. Advanced Topics, b. Recent Developments, c. Certificate Courses, d. Course Projects, e. New Software Tools, f. GATE Topics, g. NPTEL Videos, h. Swayam videos etc.

Modules	Topic / Description	Area	Remarks	Blooms Level
1	Artificial Intelligence and Artificial Neural networks	Higher studies	Gap students should be able to understand basics of Artificial Intelligence	Understand L2
2	Data warehousing	Higher studies	Gap students should be able to understand basics of Data warehousing	Understand L2
3	Data warehousing	Higher studies	Gap students should be able to understand basics of Data warehousing	Understand L2
4	DBMS	Higher studies	Gap students should be able to understand basics of DBMS	Understand L2

B. OBE PARAMETERS

1. Course Outcomes

Expected learning outcomes of the course, which will be mapped to POs. Identify a max of 2 Concepts per Module. Write 1 CO per Concept.

Modules	Course Code.#	Course Outcome At the end of the course, student should be able to . . .	Teach. Hours	Concept	Instr Method	Assessment Method	Blooms' Level
1	15CS82.1	Apply Hadoop Distributed File System for storage and processing	5	Storage, processing	Lecture	Slip Test	L3 Apply
1	15CS82.2	Map Reduce Programming to compression's.	5	Compression's, reduce	Lecture	Slip Test	L3 Apply
2	15CS82.3	Apply essential Hadoop Tools for simulation of storage and processing	5	Essentials Tools, Procedures	Lecture	Assignment	L3 Apply
2	15CS82.4	Hadoop Administration Procedures for storage and processing	5	Hadoop administration	Lecture	Assignment	L3 Apply
3	15CS82.5	Apply Business Intelligence Concepts to make decisions	5	Business intelligence	Lecture	Assignment	L3 Apply
3	15CS82.6	Application basic data for business Intelligence Concepts to make decisions	5	Data Mining	Lecture	Assignment	L3 Apply
4	15CS82.7	Apply Decision Trees Concepts to run business intelligence	5	Regression, Trees	Lecture / PPT	Assignment	L3 Apply

4	15CS82.8	Apply basic data ware housing to make a decision in business	5	corporate data organized data	Lecture / PPT	Assignment	L3 Apply
5	15CS82.9	Apply various data mining based on available text	5	Text, Web, Social	Lecture	Slip test	L3 Apply
5	15CS82.10	Apply various text mining based on available text	5	Text, Web script Social network	Lecture	Slip test	L3 Apply
-	-	Total	50	-	-	-	L2-L3

2. Course Applications

Write 1 or 2 applications per CO.

Students should be able to employ / apply the course learnings to . . .

Modules	Application Area Compiled from Module Applications.	CO	Level
1	Managing traffic on streets.	CO1	L2
1	web link-graph	CO2	L3
2	Distributed shell	CO3	L3
2	graphical user interface	CO4	L3
3	Government education,	CO5	L3
3	retailer services Image	CO6	L3
4	Processing and Character recognition	CO7	L3
4	business analysis	CO8	L3
5	Text classification/ Spam Filtering	CO9	L3
5	Sentiment Analysis	CO10	L3

3. Mapping And Justification

CO – PO Mapping with mapping Level along with justification for each CO-PO pair.

To attain competency required (as defined in POs) in a specified area and the knowledge & ability required to accomplish it.

M	Mappin	Mappi	Justification for each CO-PO pair	Le
---	--------	-------	-----------------------------------	----

Module	CO	PO	Level	Area: 'Competency' and 'Knowledge' for specified 'Accomplishment'	Level
-	CO	PO	-	'Area': 'Competency' and 'Knowledge' for specified 'Accomplishment'	-
1	CO1	PO1	3	Knowledge is required to understand Hadoop Distributed Files	L2
1	CO1	PO2	2	Analyzing the Hadoop distributed files is required to the map reduce programming.	L2
1	CO1	PO3	1	Designing a Hadoop distributed file system is required for Big data processing	L2
1	CO1	PO5	2	To store very large data sets reliably requires different tools	L3
1	CO1	PO6	3	Applying the contextual knowledge to the society using various data mining.	L3
1	CO1	PO8	1	Ethical responsibility is required to data mining using Hadoop techniques	L3
1	CO1	PO12	2	Life long learning required to sustain map reduce programming.	L3
2	CO2	PO1	3	Knowledge of Hadoop Tools is required to examine the data on a Hadoop cluster.	L3
2	CO2	PO2	1	Analyzing the data on Hadoop cluster is requires Hadoop Tools.	L3
2	CO2	PO3	2	Development of Hadoop administration requires essential Hadoop Tools.	L3
2	CO2	PO5	3	Various tools are used to manage the YARN applications.	L3
2	CO2	PO6	2	Applying the contextual knowledge to the society using different YARN applications.	L3
2	CO2	PO12	2	Life long learning required to learn essential Hadoop Tools.	L3
3	CO3	PO1	1	Knowledge is required to understand Business Intelligence Concepts and Application.	L3
3	CO3	PO2	3	Analyzing is required to understand Business Intelligence Concepts and Application	L3
3	CO3	PO3	3	Developing different Application is required Business Intelligence Concepts.	L3
3	CO3	PO5	2	Various tools are used to develop the Business applications.	L3
3	CO3	PO6	1	Applying the contextual knowledge to the society to understand Business Intelligence Concepts and Application	L3
3	CO3	PO7	2	No environment and sustainability requirement. no mapping	L3
3	CO3	PO8	2	Ethical responsibility is required to understand Business Intelligence Concepts and Applications.	L3
4	CO4	PO1	3	Knowledge of decision tree concepts are required to use the data mining techniques.	L3
4	CO4	PO2	1	Analyzing is required to reduce the association rule programming.	L3
4	CO4	PO3	1	Programs Development is required for data mining.	L3
4	CO4	PO5	3	Association rule mining tools are used to extract the knowledge in an organization.	L3
4	CO4	PO6	2	Applying the contextual knowledge to the society using various data mining.	L3
4	CO4	PO8	1	Ethical responsibility is required to data mining using Hadoop techniques	L3
4	CO4	PO12	3	Life long learning required to sustain data mining.	L3
5	CO5	PO1	1	Knowledge is required to understand hadoop Distributed Files	L3
5	CO5	PO2	2	Analysation is required to reduce the map reduce programming.	L3
5	CO5	PO3	3	Programs Development is required for data mining.	L3
5	CO5	PO5	1	Text mining tools are used to extract the knowledge in an organization.	L3
5	CO5	PO6	2	Applying the contextual knowledge to the society using various data mining.	L3
5	CO5	PO8	3	Ethical responsibility is required to data mining using hadoop techniques	L3
5	CO5	PO12	1	Life long learning required to sustain data mining.	L3

4. Articulation Matrix

CO – PO Mapping with mapping level for each CO-PO pair, with course average attainment.

-	-	Course Outcomes	Program Outcomes															-			
			At the end of the course student should be able to . . .	P	P	P	P	P	P	P	P	P	P	P	P	P	P		P	P	Le
Mo	CO.#		O	O	O	O	O	O	O	O	O	O	O	O	O	S	S	S	vel		
dul			1	2	3	4	5	6	7	8	9	1	1	1	0	1	2	1	2	3	
es																					
1	15CS 82.1	Apply Hadoop Distributed File System for storage and processing	2.5	2.5	2.5	-	2.5	2.5	-	2.5	-	-	-	2.5							L2
1	15CS 82.2	Map Reduce Programming to compression's.	2.5	2.5	2.5	-	2.5	-	-	2.5	-	-	-	2.5							L2
2	15CS 82.3	Apply essential Hadoop Tools for simulation of storage and processing	2.5	2.5	2.5	-	2.5	2.5	-		-	2.5	-	2.5							L2
2	15CS 82.4	Hadoop Administration Procedures for storage and processing	2.5	2.5	2.5	-	-	2.5	-	2.5	2.5	-	-	2.5							L3
3	15CS 82.5	Apply Business Intelligence Concepts to make decisions	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	-	-	2.5							L3
3	15CS 82.6	Application basic data for business Intelligence Concepts to make decisions	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	-	-	2.5							L3
4	15CS	Apply Decision	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	-	-	2.5							L3

	82.7	Trees Concepts to run business intelligence																			
4	15CS 82.8	Apply basic data ware housing to make a decision in business	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	2.5	-	2.5							L3
5	15CS 82.9	Apply various data mining based on available text	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	2.5	-	2.5							L3
5	15CS 82.10	Apply various text mining based on available text	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	-	-	2.5							L3
-	CS501PC	Average attainment (1, 2, or 3)	2.5	2.5	2.5	-	2.5	2.5	-	2.5	2.5	2.5	-	2.5							-
-	PO, PSO	<i>1.Engineering Knowledge; 2.Problem Analysis; 3.Design / Development of Solutions; 4.Conduct Investigations of Complex Problems; 5.Modern Tool Usage; 6.The Engineer and Society; 7.Environment and Sustainability; 8.Ethics; 9.Individual and Teamwork; 10.Communication; 11.Project Management and Finance; 12.Life-long Learning; S1.Software Engineering; S2.Data Base Management; S3.Web Design</i>																			

5. Curricular Gap and Content

Topics & contents not covered (from A.4), but essential for the course to address POs and PSOs.

Modules	Gap Topic	Actions Planned	Schedule Planned	Resources Person	PO Mapping
1	Artificial intelligence	12/02/19	12/02/19	self	p02
2	dbms	05/03/19	05/03/19	self	p02
3	Data warehousing	10/04/19	10/04/19	self	p02

6. Content Beyond Syllabus

Topics & contents required (from A.5) not addressed, but help students for Placement, GATE, Higher Education, Entrepreneurship, etc.

Mo	Gap Topic	Area	Actions	Schedule	Resources	PO
----	-----------	------	---------	----------	-----------	----

Modules			Planned	Planned	Person	Mapping
1	Essential tools	28/03/19	28/03/19	self	p02	p02
2	Essential tools	28/03/19	01/04/19	self	p02	p02
3	Essential tools	28/03/19	01/04/19	self	p02	p02
4	Corporate Data	10/04/19	10/04/19	self	p02	p02
5	Text Mining	11/04/19	11/04/19	self	p02	p02

C. COURSE ASSESSMENT

1. Course Coverage

Assessment of learning outcomes for Internal and end semester evaluation.

Distinct assignment for each student. 1 Assignment per chapter per student. 1 seminar per test per student.

Modules	Title	Teaching Hours	No. of question in Exam						CO	Levels
			CIA-1	CIA-2	CIA-3	Asg	Extra Asg	SEE		
1	Hadoop Distributed File System Basics	10	2	-	-	1	1	1	CO1,CO2	L2,L3
2	Essential Hadoop Tools	10	2	-	-	1	1	1	CO3,CO4	L3
3	Basic Intelligence Concepts and Applications	10	-	2	-	1	1	1	CO5,CO6	L3
4	Decision Trees	10	-	2	-	1	1	1	CO7,CO8	L3
5	Text Mining	10	-	-	4	1	1	1	CO9,CO10	L3
-	Total	50	4	4	4	5	5	5	-	-

2. Continuous Internal Assessment (CIA)

Assessment of learning outcomes for Internal exams. Blooms Level in last column shall match with A.2.

Modules	Evaluation	Weightage in Marks	CO	Levels
---------	------------	--------------------	----	--------

1, 2	CIA Exam - 1	15	CO1, CO2, CO3, CO4	L2,L3
3, 4	CIA Exam - 2	15	CO5,CO6,CO7,CO8	L3
5	CIA Exam - 3	15	CO9,CO10,	L3
1, 2	Assignment - 1	05	CO1, CO2, CO3, CO4	L3
3, 4	Assignment - 2	05	CO5,CO6,CO7,CO8	L3
5	Assignment - 3	05	CO9,CO10,	L3
1, 2	Seminar - 1	-		
3, 4	Seminar - 2	-		
5	Seminar - 3	-		
1, 2	Quiz - 1		CO1 to CO10	L3
3, 4	Quiz - 2	20	-	
5	Quiz - 3		-	-
1 - 5	Other Activities - Mini Project	-	CO9, CO10	L2,L2
	Final CIA Marks	20	-	-

D1. TEACHING PLAN - 1

Module - 1

Title:	Hadoop Distributed File System Basics,	Appr Tim e:	10 Hrs
a	Course Outcomes	CO	Blooms
-	At the end of the topic the student should be able to . . .	-	Level
1	Apply hadoop Distributed File System Basics to MapReduce Programming	CO1	L3

2	Apply Map Reduce Programming to compression's.	CO2	L3
b	Course Schedule	-	-
Class No	Portion covered per hour	-	-
1	Hadoop Distributed File System Basics,	CO1	L3
2	Hadoop Distributed File System Basics,	CO1	L3
3	Hadoop Distributed File System Basics,	CO1	L3
4	Hadoop Distributed File System Basics,	CO1	L3
5	Running Example Programs and Benchmarks,,	CO2	L3
6	Running Example Programs and Benchmarks,,	CO2	L3
7	Running Example Programs and Benchmarks,,	CO2	L3
8	Hadoop Map Reduce Framework	CO2	L3
9	Map Reduce Programming	CO2	L3
10	Map Reduce Programming	CO2	L3
c	Application Areas	-	-
-	Students should be able employ / apply the Module learnings to . . .	-	-
1	Managing traffic on streets.	CO1	L2
2	web link-graph	CO2	L3
d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
1	What are the 3 Vs of Big Data?	CO1	L2
2	How does Big Data impact the business models?	CO1	L2
3	What is Hadoop?	CO1	L2
4	How does Map-Reduce algorithm work?	CO1	L3
5	What are the key issues in managing Big Data?	CO1	L2
6	What is Hadoop? Name the Main Components of a Hadoop Application.	CO1	L3
7	What do you understand by "Rack Awareness"?	CO1	L3
8	What is Speculative Execution?	CO1	L3
9	State some of the important features of Hadoop.	CO2	L3
10	How can you differentiate RDBMS and Hadoop?	CO2	L3
11	What are active and passive Name Nodes?	CO2	L3
12	What are the Components of Apache HBase?	CO2	L3
13	How is the Data Node failure handled by Name Node?	CO2	L3
14	Explain the Name Node recovery process.	CO2	L3
15	What are the basic steps involved in map reduce data flow?	CO2	L3
e	Experiences	-	-
1		CO1	L2
2			
3			
4		CO2	L3
5			

Module – 2

Title:	Essential Hadoop Tools	Appr Tim	10 Hrs
---------------	------------------------	---------------------	-------------------

		e:	
a	Course Outcomes	CO	Blooms Level
-	At the end of the topic the student should be able to . . .	-	
1	Apply essential Hadoop Tools for simulation of storage and processing	CO3	L3
2	Apply Hadoop Administration Procedures for storage and processing	CO4	L3
b	Course Schedule	-	-
Class No	Portion covered per hour	-	-
11	Essential Hadoop Tools	CO3	L2
12	Essential Hadoop Tools	CO3	L3
13	Hadoop YARN Applications	CO3	L3
14	Hadoop YARN Applications	CO3	L3
15	Hadoop YARN Applications	CO3	L3
16	Managing Hadoop with Apache Ambari	CO4	L3
17	Managing Hadoop with Apache Ambari	CO4	L3
18	Basic Hadoop Administration Procedures	CO4	L3
19	Basic Hadoop Administration Procedures	CO4	L3
20	Basic Hadoop Administration Procedures	CO4	L3
c	Application Areas	-	-
-	Students should be able employ / apply the Module learnings to . . .	-	-
1	Distributed shell	CO3	L3
2	graphical user interface	CO4	L3
d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
16	What are the main components of Job flow in YARN architecture ?	CO3	L2
17	What is the role of Application Master in YARN architecture ?	CO3	L2
18	Write the structure of YARN applications.	CO3	L2
19	Write a Apache Ambari dashboard view of hadoop cluster.	CO4	L3
20	What are the different views of Apache Ambari.	CO4	L3
21	Write the basicHDFS administration.	CO4	L3
22	Explain capacity scheduler background.	CO4	L3
e	Experiences	-	-
1		CO3	L2
2			
3			
4		CO4	L3
5			

E1. CIA EXAM – 1

a. Model Question Paper - 1

Crs Code:	15CS82	Sem:	VIII	Marks:	30	Time:	75 minutes	
Course:	BIG DATA ANALYTICS							
-	-	Note: Answer all questions, each carry equal marks. Module : 1, 2				Marks	CO	Level
1	a	What do you understand by "Rack Awareness"?				8	co1	L3
	b	What is Hadoop? Name the Main Components of a Hadoop Application				7	co1	L3

		OR			
1	a	What are the basic steps involved in map reduce data flow?	8	CO2	L3
	b	Explain HDFS data storage.	7	CO2	L3
2	a	What are the different views of Apache Ambari.	8	CO3	L2
	b	List and Describe essential HADOOP Tools.	7	CO3	L3
		OR			
2	a	Write the structure of YARN applications	8	CO4	L2
	b	Explain YARN Application Framework.	7	CO4	L3

b. Assignment -1

Note: A distinct assignment to be assigned to each student.

Model Assignment Questions							
Crs Code:	15CS82	Sem:	Viii	Marks:	5	Time:	90 – 120 minutes
Course:	BIG DATA ANALYTICS			Module :	1, 2		
Note: Each student to answer 2-3 assignments. Each assignment carries equal mark.							
SNo	USN	Assignment Description	Marks	CO	Level		
1		State some of the important features of Hadoop.	5	CO1	L3		
2		How does Big Data impact the business models?	5	CO1	L3		
3		What are active and passive Name Nodes?	4	CO1	L3		
4		How does Map-Reduce algorithm work?	5	CO2	L3		
5		What are the key issues in managing Big Data?	8	CO2	L3		
6		How does Map-Reduce algorithm work?	9	CO2	L3		
7		What are the key issues in managing Big Data?	6	CO2	L2		
8		What is Hadoop? Name the Main Components of a Hadoop Application.	9	CO2	L3		
9		What do you understand by "Rack Awareness"?	8	CO2	L3		
10		What is Speculative Execution?	6	CO2	L3		
11		How is the Data Node failure handled by Name Node?	9	CO3	L3		
12		What is the role of Application Master in YARN architecture ?	10	CO3	L2		
13		Write the structure of YARN applications.	7	CO4	L2		
14		Write a Apache Ambari dashboard view of hadoop cluster.	8	CO4	L3		
15		State some of the important features of Hadoop.	5	CO1	L3		
16		How does Big Data impact the business models?	5	CO1	L3		
17		What are active and passive Name Nodes?	4	CO1	L3		
18		How does Map-Reduce algorithm work?	5	CO2	L3		
19		What are the key issues in managing Big Data?	8	CO2	L3		
20		How does Map-Reduce algorithm work?	9	CO2	L3		
21		What are the key issues in managing Big Data?	6	CO2	L2		
22		What is Hadoop? Name the Main Components of a Hadoop Application.	9	CO2	L3		

23		What do you understand by "Rack Awareness"?	8	CO2	L3
24		What is Speculative Execution?	6	CO2	L3
25		How is the Data Node failure handled by Name Node?	9	CO3	L3
26		What is the role of Application Master in YARN architecture ?	10	CO3	L2

D2. TEACHING PLAN - 2

Module - 3

Title:	Business Intelligence Concepts and Applications	App r Tim e:	12 Hrs
a	Course Outcomes	CO	Blooms Level
-	At the end of the topic the student should be able to . . .	-	-
1	Apply Business Intelligence Concepts to make decisions	CO5	L3
2	Application basic data for business Intelligence Concepts to make decisions	CO6	L3
b	Course Schedule		
Class No	Portion covered per hour	-	-
21	Business Intelligence Concepts and Application	CO5	L2
22	Business Intelligence Concepts and Application	CO5	L3
23	Business Intelligence Concepts and Application	CO5	L3
24	Business Intelligence Concepts and Application	CO5	L3
25	Data Warehousing	CO6	L3
26	Data Warehousing	CO6	L3
27	Data Mining	CO6	L3
28	Data Mining	CO6	L3
29	Data Visualization	CO6	L3
30	Data Visualization	CO6	L3
c	Application Areas	-	-
-	Students should be able employ / apply the Module learnings to . . .	-	-
1	Government education,	CO5	L3
2	retailer services Image	CO6	L3
d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
23	Describe the Business Intelligence and Data Mining cycle.	CO5	L2
24	Describe the data processing chain.	CO5	L3
25	What are the similarities between diamond mining and data mining?	CO5	L2
26	What are the different data mining techniques? Which of these would be relevant in your current work?	CO5	L3
27	What is a dashboard? How does it help?	CO5	L2
28	Create a visual to show the weather pattern in your city. Could you show together temperature, humidity, wind, and rain/snow over a period of time.	CO5	L3
29	Why should organizations invest in business intelligence solutions? Are these more important than IT security solutions? Why or why not?	CO5	L3
30	List 3 business intelligence applications in the hospitality industry.	CO5	L3
31	Describe 2 BI tools used in your organization.	CO5	L2
32	Businesses need a 'two-second advantage' to succeed. What does that mean	CO5	L3

	to you?		
33	What is the purpose of a data warehouse?	CO6	L2
34	What are the key elements of a data warehouse? Describe each one.	CO6	L3
35	What are the sources and types of data for a data warehouse?	CO6	L3
36	How will data warehousing evolve in the age of social media?	CO6	L3
37	What is data mining? What are supervised and unsupervised learning techniques?	CO6	L2
38	Describe the key steps in the data mining process. Why is it important to follow these processes?	CO6	L3
39	What are the major mistakes to be avoided when doing data mining?	CO6	L3
40	What are the key requirements for a skilled data analyst?	CO6	L2
41	What are some of the most popular data mining techniques?	CO6	L3
42	What is a confusion matrix?	CO6	L2
43	Why is data preparation so important and time consuming?	CO6	L3
44	What is data visualization?	CO6	L3
45	What are some key requirements for good visualization.	CO6	L2
46	Describe some key steps in data visualization.	CO6	L3
47	What are the data visualization techniques? When would you use table or graphs?	CO6	L3
48	How would you judge the quality of data visualizations?	CO6	L3
e	Experiences	-	-
1		CO6	L2
2			
3			
4		CO6	L3
5			

Module - 4

Title:	Decision Trees	App r Tim e:	10 Hrs
a	Course Outcomes	CO	Blooms Level
-	At the end of the topic the student should be able to . . .	-	
1	Apply Decision Trees Concepts to run business intelligence	CO7	L3
2	Apply basic data ware housing to make a decision in business	CO8	L3
b	Course Schedule		
Class No	Portion covered per hour	-	-
31	Decision Trees	CO7	L2
32	Decision Trees	CO7	L3
33	Regression	CO7	L3
34	Regression	CO7	L3
35	Artificial Neural Networks	CO8	L3
36	Artificial Neural Networks	CO8	L3
37	Cluster Analysis	CO8	L3
38	Cluster Analysis	CO8	L3
39	Association Rule Mining	CO8	L3
40	Association Rule Mining	CO8	L3

c	Application Areas	-	-
-	Students should be able employ / apply the Module learnings to ...	-	-
1	Processing and Character recognition	CO7	L3
2	business analysis	CO8	L3
d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
49	What is a decision tree? Why are decision trees the most popular classification technique?	CO7	L2
50	What is a splitting variable? Describe three criteria for choosing splitting variable.	CO7	L3
51	What is pruning? What are pre-pruning and post-pruning? Why choose one over the other?	CO7	L3
52	What are gini's coefficient, and information gain?	CO7	L3
53	What is a regression model?	CO8	L2
54	What is a scatter plot? How does it help?	CO8	L3
55	Compare and contrast decision trees with regression models?	CO8	L2
56	What is a neural network? How does it work?	CO8	L3
57	Compare a neural network with a decision tree.	CO8	L2
58	What makes a neural network versatile enough for supervised as well as non-supervised learning tasks?	CO8	L3
59	Examine the steps in developing a neural network for predicting stock prices.	CO8	L3
60	What is unsupervised learning? When is it used? Describe three business applications in your industry where cluster analysis will be useful.	CO8	L3
61	What are association rules? How do they help?	CO8	L3
62	How many association rules should be used?	CO8	L2
e	Experiences	-	-
1		CO7	L2
2			
3			
4		CO8	L3
5			

E2. CIA EXAM – 2

a. Model Question Paper - 2

Crs Code:	15CS82	Sem:	VIII	Marks:	30	Time:	75 minutes	
Course:	BIG DATA ANALYTICS							
-	-	Note: Answer all questions, each carry equal marks. Module : 3, 4				Marks	CO	Level
1	a	List 3 business intelligence applications in the hospitality industry.	5	CO	L2	5		
	b	What is the purpose of a data warehouse?	4	CO	L2	5		
	c	What are the sources and types of data for a data warehouse?	6	CO	L3	5		
OR								
1	a	How will data warehousing evolve in the age of social media?	8	CO	L3	6		
	b	What is data mining? What are supervised and unsupervised learning techniques?	7	CO	L3			

				6	
2	a	What is a decision tree? Why are decision trees the most popular classification technique?	7	CO 7	L3
	b	What is a splitting variable? Describe three criteria for choosing splitting variable.	8	CO 7	L3
		OR			
2	a	What is a regression model?	5	CO 8	L2
	b	What is a scatter plot? How does it help?	5	CO 8	L3
		What is a neural network? How does it work?	5	CO 8	L3

b. Assignment – 2

Note: A distinct assignment to be assigned to each student.

Model Assignment Questions							
Crs Code:	15CS82	Sem:	VIII	Marks:	5	Time:	90 – 120 minutes
Course:	BIG DATA ANALYTICS			Module :	3, 4		
Note: Each student to answer 2-3 assignments. Each assignment carries equal mark.							
SNo	USN	Assignment Description			Marks	CO	Level
1		List 3 business intelligence applications in the hospitality industry.			5	CO5	L3
2		Describe 2 BI tools used in your organization.			8	CO5	L2
3		Businesses need a 'two-second advantage' to succeed. What does that mean to you?			8	CO5	L3
4		What is the purpose of a data warehouse?			4	CO6	L2
5		What are the key elements of a data warehouse? Describe each one.			8	CO6	L3
6		What are the sources and types of data for a data warehouse?			6	CO6	L3
7		How will data warehousing evolve in the age of social media?			9	CO6	L3
8		What is data mining? What are supervised and unsupervised learning techniques?			8	CO6	L2
9		What is a decision tree? Why are decision trees the most popular classification technique?			7	CO7	L2
10		What is a splitting variable? Describe three criteria for choosing splitting variable.			8	CO7	L3
11		What is pruning? What are pre-pruning and post-pruning? Why choose one over the other?			8	CO7	L3
12		What are gini's coefficient, and information gain?			8	CO7	L4
13		What is a regression model?			5	CO8	L2

14	What is a scatter plot? How does it help?	5	CO8	L3
15	Compare and contrast decision trees with regression models?	8	CO8	L2
16	What is a neural network? How does it work?	7	CO8	L3
17	Compare a neural network with a decision tree.	7	CO8	L2
18	What makes a neural network versatile enough for supervised as well as non-supervised learning tasks?	8	CO8	L3
19	List 3 business intelligence applications in the hospitality industry.	5	CO5	L3
20	Describe 2 BI tools used in your organization.	8	CO5	L2
21	Businesses need a 'two-second advantage' to succeed. What does that mean to you?	8	CO5	L3
22	What is the purpose of a data warehouse?	4	CO6	L2
23	What are the key elements of a data warehouse? Describe each one.	8	CO6	L3
24	What are the sources and types of data for a data warehouse?	6	CO6	L3
25	How will data warehousing evolve in the age of social media?	9	CO6	L3
26	What is data mining? What are supervised and unsupervised learning techniques?	8	CO6	L2

D3. TEACHING PLAN - 3

Module - 5

Title:	Text Mining	App r Tim e:	10 Hrs
a	Course Outcomes	CO	Blooms Level
-	At the end of the topic the student should be able to . . .	-	
1	Apply various data mining techniques based on available text	CO9	L3
2	Apply various text mining techniques based on available text	CO10	L3
b	Course Schedule	-	-
Class No	Portion covered per hour	-	-
41	Text Mining	CO9	L2
42	Text Mining	CO9	L3
43	Naïve-Bayes Analysis	CO9	L3
44	Naïve-Bayes Analysis	CO9	L3
45	Support Vector Machines	CO10	L3

46	Support Vector Machines	CO10	L3
47	Web Mining	CO10	L2
48	Web Mining	CO10	L3
49	Social Network Analysis	CO10	L3
50	Social Network Analysis	CO10	L3
c	Application Areas	-	-
-	Students should be able employ / apply the Module learnings to ...	-	-
1	Text classification/ Spam Filtering	CO9	L3
2	Sentiment Analysis	CO10	L3
d	Review Questions	-	-
-	The attainment of the module learning assessed through following questions	-	-
63	Why is text mining useful in the age of social media?	CO9	L3
64	What kinds of problems can be addressed using text mining?	CO9	L3
65	What kinds of sentiments can be found in the text?	CO9	L3
66	What are the three types of web mining?	CO10	L3
67	What are the two major ways that a website can become popular?	CO10	L3
68	What are the privacy issues in web mining?	CO10	L3
69	A user spends 60 minutes on the web, visiting 10 web pages in all. Given the click stream data, what kind of an analysis would you do?	CO10	L3
70	What is click stream analysis?	CO10	L3
e	Experiences	-	-
1		CO10	L2
2		CO9	
3			
4		CO9	L3
5			

E3. CIA EXAM – 3

a. Model Question Paper - 3

Crs Code:	15CS82	Sem:	VIII	Marks:	30	Time:	75 minutes	
Course:	BIG DATA ANALYTICS							
-	-	Note: Answer all questions, each carry equal marks. Module : 5				Marks	CO	Level
1	a	Why is text mining useful in the age of social media?				8	CO 9	L3
	b	What kinds of problems can be addressed using text mining?				7	CO 9	L3
	c							
OR								
1	a	Briefly explain support vector machine.				8	CO 9	L3
	b	Briefly explain naive-Bayes analysis				9	CO 9	L3
	c							
2	a	What are the three types of web mining?				6	CO1	L3

				0	
	b	What kinds of sentiments can be found in the text?	5	CO1 0	L3
	c				
OR					
2	a	What are the two major ways that a website can become popular?	8	CO1 0	L3
	b	What are the privacy issues in web mining?	5	CO1 0	L3
	c	What is click stream analysis?	2	CO1 0	L2

b. Assignment – 3

Note: A distinct assignment to be assigned to each student.

Model Assignment Questions							
Crs Code:	15CS82	Sem:	VIII	Marks:	5	Time:	90 – 120 minutes
Course:	BIG DATA ANALYTICS			Module :	5		
Note: Each student to answer 2-3 assignments. Each assignment carries equal mark.							
SNo	USN	Assignment Description	Marks	CO	Level		
1		Why is text mining useful in the age of social media?	8	CO9	L3		
2		What kinds of problems can be addressed using text mining?	7	CO9	L3		
3		What are the two major ways that a website can become popular?	8	CO9	L3		
4		What are the privacy issues in web mining?	5	CO9	L3		
5		What is click stream analysis?	2	CO1 0	L2		
6		Briefly explain support vector machine.	8	CO1 0	L3		
7		Briefly explain naive-Bayes analysis	9	CO1 0	L3		
8		What are the three types of web mining?	6	CO9	L3		
9		What kinds of sentiments can be found in the text?	5	CO1 0	L3		
10		Why is text mining useful in the age of social media?	8	CO9	L3		
11		What kinds of problems can be addressed using text mining?	7	CO9	L3		
12		What are the two major ways that a website can become popular?	8	CO9	L3		
13		What are the privacy issues in web mining?	5	CO9	L3		
14		What is click stream analysis?	2	CO1	L2		

				0	
15		Briefly explain support vector machine.	8	CO1 0	L3
16		Briefly explain naive-Bayes analysis	9	CO1 0	L3
17		What are the three types of web mining?	6	CO9	L3
18		What kinds of sentiments can be found in the text?	5	CO1 0	L3
19		Why is text mining useful in the age of social media?	8	CO9	L3
20		What kinds of problems can be addressed using text mining?	7	CO9	L3
21		What are the two major ways that a website can become popular?	8	CO9	L3
22		What are the privacy issues in web mining?	5	CO9	L3
23		What is click stream analysis?	2	CO1 0	L2
24		Briefly explain support vector machine.	8	CO1 0	L3
25		Briefly explain naive-Bayes analysis	9	CO1 0	L3
26		What are the three types of web mining?	6	CO9	L3

F. EXAM PREPARATION

1. University Model Question Paper

Course:	BIG DATA ANALYTICS				Month / Year	May /2018		
Crs Code:	15CS82	Sem:	VIII	Marks:	80	Time:	180 minutes	
Mod ule	Note	Answer all FIVE full questions. All questions carry equal marks.				Marks	CO	Level
1	a							
	b							
	c							
	d							
		OR						
1	a							
	b							
	c							
	d							
2	a							
	b							
	c							
	d							
		OR						
2	a							
	b							
	c							
	d							

3	a				
	b				
	c				
	d				
OR					
3	a				
	b				
	c				
	d				
4	a				
	b				
	c				
	d				
OR					
4	a				
	b				
	c				
	d				
5	a				
	b				
	c				
	d				
OR					
5	a				
	b				
	c				
	d				

2. SEE Important Questions

Course:	BIG DATA ANALYTICS				Month / Year	May /2018	
Crs Code:	15CS82	Sem:	8	Marks:	80	Time:	180 minutes
	Note Answer all FIVE full questions. All questions carry equal marks.					-	-
Module	Qno.	Important Question	Marks	CO	Year		
1	a	What are active and passive NameNodes?	6	CO1	L3		
	b	How does Map-Reduce algorithm work?	8	CO1	L3		
	c	What are the key issues in managing Big Data?	6	CO1	L3		
	d	What do you understand by "BLOCK Replication"?	8	CO2	L2		
	e	What is Hadoop? Name the Main Components of a Hadoop Application.	7	CO2	L3		
	f	What do you understand by "Rack Awareness"?	5	CO2	L3		
2	a	What are the main components of Job flow in YARN architecture ?	10	CO3	L2		
	b	What is the role of Application Master in YARN architecture ?	6	CO3	L2		
	c	Write the structure of YARN applications.	4	CO3	L2		
	d	Write a Apache Ambari dashboard view of hadoop cluster.	7	CO4	L3		
	e	What are the different views of Apache Ambari.	3	CO4	L3		
	f	Write the basic HDFS administration.	6	CO4	L3		
3	a	Describe the Business Intelligence and Data Mining cycle.	7	CO5	L2		
	b	What are the different data mining techniques? Which of these would be relevant in your current work?	6	CO5	L3		
	c	What are the similarities between diamond mining and data mining?	3	CO5	L2		
	d	What is a dashboard? How does it help?	4	CO6	L2		
	e	Create a visual to show the weather pattern in your city. Could you show together temperature, humidity, wind, and rain/snow over a period of time.	8	CO6	L3		

	f	Why should organizations invest in business intelligence solutions? Are these more important than IT security solutions? Why or why not?	4	CO6	L3
4	a	What is a decision tree? Why are decision trees the most popular classification technique?	8	CO7	L2
	b	What is a splitting variable? Describe three criteria for choosing splitting variable.	8	CO7	L3
	c	What is a regression model?	2	CO8	L2
	d	Compare and contrast decision trees with regression models?	6	CO8	L2
	e	What is a neural network? How does it work?	8	CO8	L3
5	a	Why is text mining useful in the age of social media?	5	CO9	L3
	b	What kinds of problems can be addressed using text mining?	6	CO9	L3
	c	What kinds of sentiments can be found in the text?	5	CO9	L3
	d	What are the three types of web mining?	5	CO10	L3
	e	What are the two major ways that a website can become popular?	6	CO10	L3
		What are the privacy issues in web mining?	5	CO10	L3

G. Content to Course Outcomes

1. TLPA Parameters

Table 1: TLPA – Example Course

Module-#	Course Content or Syllabus (Split module content into 2 parts which have similar concepts)	Content Teaching Hours	Blooms' Learning Levels for Content	Final Bloom Levels	Identified Action Verbs for Learning	Instruction Methods for Learning	Assessment Methods to Measure Learning
A	B	C	D	E	F	G	H
1	Hadoop Distributed File System Basics, Running Example Programs and Benchmarks, Hadoop MapReduce Framework, Map Reduce Programming	5	- L1 - L2 - L3	L3	- Understand - Apply	- Lecture	- Slip Test
1	Essential Hadoop Tools, Hadoop YARN Applications, Managing Hadoop with Apache Ambari, Basic Hadoop Administration Procedures	5	- L3	L3	- Apply	- Lecture	- Assignment
2	Business Intelligence Concepts and Application, Data Warehousing, Data Mining, Data Visualization	5	- L2 - L3	L3	- Apply	- Lecture	- Assignment

2	Decision Trees, Regression, Artificial Neural Networks, Cluster Analysis, Association Rule Mining	5	- L2 - L3	L3	-Apply	- Lecture	- Slip Test
3	Text Mining, Naïve-Bayes Analysis, Support Vector Machines, Web Mining, Social Network Analysis	5	- L2 - L3	L3	-Apply	- Lecture	- Slip Test
3	Test Execution: Overview of test execution, from test case specification to test cases, Scaffolding, Generic versus specific scaffolding, Test oracles, Self-checks as oracles, Capture and replay	5	- L2 - L3	L3	-Apply	- Lecture	- Assignment
4	Process Framework :Basic principles: Sensitivity, redundancy, restriction, partition, visibility, Feedback, the quality process, Planning and monitoring, Quality goals, Dependability properties, Analysis Testing, Improving the process, Organizational factors. Planning and Monitoring the Process: Quality and process, Test and analysis strategies and plans, Risk planning, monitoring the process, Improving the 10 Hours process, the quality team.	5	- L2 - L3	L3	-Apply	- Lecture	- Assignment
4	Documenting Analysis and Test: Organizing documents, Test strategy document, Analysis and test plan, Test design specifications documents, Test and analysis reports	5	- L2 - L3	L3	-Apply	- Lecture	- Assignment
5	Integration and Component-Based Software Testing: Overview, Integration testing strategies, Testing components and assemblies. System, Acceptance and Regression Testing: Overview, System testing, Acceptance testing, Usability, Regression testing, Regression test selection techniques, Test case prioritization and selective execution.	5	- L2 - L3	L3	-Apply	- Lecture	- Assignment
5	Levels of Testing, Integration Testing: Traditional view of testing levels, Alternative life-cycle models, The SATM system, Separating integration and system testing, A closer look at the SATM system, Decomposition-based, call graph-based, Path-based integrations.	5	- L2 - L3	L3	-Apply	- Lecture	- Assignment

2. Concepts and Outcomes:

Table 2: Concept to Outcome – Example Course

M	Learning	Identified	Final	Concept	CO	Course
od	or	Conce	Concept	Justification	Components	Outcome
e-	from	pts		(What all Learning	(1.Action Verb,	

#	study of the Content or Syllabus	from Content		Happened from the study of Content / Syllabus. A short word for learning or outcome)	2.Knowledge, 3.Condition / Methodology, 4.Benchmark)	Student Should be able to ...
A	I	J	K	L	M	N
1	- Hadoop Distributed File System Basics, - Running Example Programs and Benchmarks	- Storing data , - processing data	Storage processing	Provides reliable and scalable distributed computing	- Apply - large data - hadoop distributed file system	Apply Hadoop Distributed File System for storage and processing
1	-Hadoop MapReduce Framework, - Map Reduce Programming	- Mapper - reducer	Map reduce	Allows processing a large data set	- Apply - map reduce - structured data	Map Reduce Programming to compression's.
2	-Essential Hadoop Tools, - Hadoop YARN Applications,	- Essentials Tools, - Procedures	Hadoop Tools,	Used to support hadoop distributed file system	- Apply - applications - HDFS	Apply essential Hadoop Tools for simulation of storage and processing
2	- Managing Hadoop with Apache Ambari, - Basic Hadoop Administration Procedures	- Hadoop administrator	- Hadoop Administration	Manage and monitor the hadoop clusters	- Apply - Hadoop clusters - administrations	Hadoop Administration Procedures for storage

		tion - Hadoop management				and processing
3	-Business Intelligence Concepts and Application,	- Business concepts - Business concept applications	Business intelligence	Supports decision making for variety of problems	- Apply - business knowledge - business intelligence	Apply Business Intelligence Concepts to make decisions
3	-Data Warehousing, - Data Mining, - Data Visualization	- storage - extract - visualize	Data mining	Discovers the knowledge, insights and patterns	- Apply - knowledgeable data -data mining	Application basic data for business Intelligence Concepts to make decisions
4	-Decision Trees, -Regression,	Regression, Trees	Regression equation	Helps classification.	In- Apply - classified data - regression model	Apply Decision Trees Concepts to run

						business intelligence
4	-Artificial Neural Networks, - Cluster Analysis, - Association Rule Mining	corporate data organized data	organized data	discover affinities between products in transactions	- Apply organized data - association rules	Apply basic data warehousing to make a decision in business
5	-Text Mining, -Naïve-Bayes Analysis,	Text, Navie - Bayes	Text Mining	discover valuable insights about the business	- Apply data insights - text mining	Apply various data mining based on available text
5	-Support Vector Machines, - Web Mining, - Social Network Analysis	Text, Web script Social network	Web mining	Build the authority for the web sites	- Apply website authority - web mining	Apply various text mining based on available text