

GANPAT UNIVERSITY

FACULTY OF TECHNOLOGY

Programme	Bachelor of Technology				Branch/Spec.	Computer Science & Engineering (BDA)			
Semester	VII				Version	1.0.1.0			
Effective from Academic Year		2018 – 19			Effective for the batch Admitted in			June 2016	
Subject code	2CSE705		Subject Name		BIG DATA APPLICATION DEVELOPMENT				
Teaching scheme					Examination scheme (Marks)				
(Per week)	Lecture(DT)		Practical(Lab.)		Total		CE	SEE	Total
	L	TU	P	TW					
Credit	0	0	1	0	1	Theory	-	-	-
Hours	0	0	2	0	2	Practical	40	60	100
Pre-requisites:									
Knowledge of Programming Languages, Software Engineering									
Learning Outcome:									
After successful completion of the course students should be able to									
<ul style="list-style-type: none"> • Design architecture of big data applications • Develop big data based applications • Use publicly available distributions for big data • To analyze and trouble shoot the problems while developing application on big data • Deploy the application based on real life 									
Content									
<p>This module is focused on developing applications in the big data. They have to follow below topics to develop the application.</p> <ul style="list-style-type: none"> • Big data based applications : Understanding distributions and ecosystems for big data,working with HDFS • Designing Code for The big data :MapReduce, Hive, Pig, hadoop • Frameworks: Working with different ecosystems, working with different file formats: JSON, CSV,TSV. Working with flume, hive ,zookeeper, avro,jaql, Sqoop • Working with oozie to perform scheduling of jobs. <p>By the end of this module the student will have a detailed overview of the design and development process involved in creating a Big Data application.</p>									
Reference Books									

1	Big Data, Black Book: Covers Hadoop ,MapReduce, Hive, YARN, Pig, R and Data Visualization, DT Editorial Services.
2	Tom White, Hadoop: The Definitive Guide, O'Reilly Media, Third Edition, 2012.
3	Bill Franks, Taming The Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics, Wiley, 2012.