|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **NAME OF DEPARTMENT:** | | | | | | | | | | | | | | | | | | School of Computing | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **Course Name:** | | | | | | | | **Bachelor of Computer Applications (BCA)** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  | | | | |
|  | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | |  | | | | | | | |  | | | | |
| **Subject Name:** | | | | | | | | Fundamental of Data Analytics | | | | | | | | | | | | | | | | | | | | | | | | **Subject Code:** | | | | | | | | TBC604(1) | | | | |
|  | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | |  | | | | | | | | |  | | | | |
| **1** | **Contact Hours:** | | | | | | | | | 48 | | | | |  | | | | | | | | | | | | | | | | | | **L** | | 3 | | | | **T** | | | 1 | **P** | 0 | | |
|  |  | | | | | | | | |  | | | | |  | | | | | | | | | | | | | | | | | |  | |  | | | |  | | |  |  |  | | |
| **2** | **Examination Duration (Hrs):** | | | | | | | | | | | | | | | | | | | |  | **Theory** | | | | | 0 | 3 |  | **Practical** | | | | 0 | | | 0 | |  | | | | | |
|  |  | | | | | | | | | | | | | | | | | | | |  |  | | | | |  |  |  |  | | | |  | | |  | |  | | | | | |
| **3** | **Relative Weightage:** | | | | | | | | | | | |  | | | | | **CWE:** | | | | | | | 25 | | **MTE:** | | | 25 | | **ETE:** | | | | 50 | | | | |  | | | | |
|  |  | | | | | | | | | | | |  | | | | |  | | | | | |  | | |  | | |  | |  | | | |  | | | | |  | | | | |
| **4** | **Credits:** | | | | | 0 | | | 4 |  | | | | | | | | | | | | | |  | | |  | | |  | |  | | | |  | | | | |  | | | | |
|  |  | | | | |  | | |  |  | | | | | | | | | | | | | |  | | |  | | |  | |  | | | |  | | | | |  | | | | |
| **5** | **Semester:** | | | | | |  | | |  | | | | **\*** | | |  | | |  | | |  | | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | **Autumn** | | | | | | **Spring** | | | | | | | | **Both** | | | | | | |  | | | | | | | | | | | | | | | | | | |
|  |  | | | |  | | | | | |  | | | | | | | |  | | | | | | |  | | | | | | | | | | | | | | | | | | |
| **6** | **Pre-Requisite:** | | | | | | | | | | | Fundamentals of computer science | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **7** | **Subject Area:** | | | | | | | | | | | Computer Science | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **8** | **Objective:** | | | | | | | | | | | To teach data science concepts and applications of Python programming. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **9** | **Course Outcome:** | | | | | | | | | | | A student who successfully fulfills the course requirements will be able to: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **CO 1** | | | | | | | | | | | Understand the basics of data science. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **CO 2** | | | | | | | | | | | Knowledge of advanced analytics technology and tools. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **CO 3** | | | | | | | | | | | Knowledge of Hadoop distributed file system architecture. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **CO 4** | | | | | | | | | | | Processing the data with MapReduce. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **CO 5** | | | | | | | | | | | Application of MapReduce | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **10** | | **Details of the Course:** | | | | | | | | | | | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **Unit No.** | | | | **CONTENT** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | **CONTACT HOURS** | | | | | | |
| **1** | | | | INTRODUCTION TO DATA SCIENCE: Introduction: Overviews of Big Data, State of the Practice in Analytics, The Data Scientist, Big Data Analytics in Industry Verticals, Data Analytics Lifecycle Challenges of Conventional Systems, Statistical Concepts: Sampling Distributions, Re-Sampling, Statistical Inference, Prediction Error, Regression Modelling, Multivariate Analysis , Bayesian Modelling. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 8 | | | | | | |
| **2** | | | | NoSQL Databases: - RDBMS Vs NoSQL, Types of No SQL Databases, Architecture of NoSQL Databases, CAP Theorem, HBase Architecture, Reading and writing data | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 10 | | | | | | |
| **3** | | | | ADVANCED ANALYTICS TECHNOLOGY AND TOOLS. Analytics for Unstructured Data, Use Cases, MapReduce, Apache Hadoop, The Hadoop Ecosystem, Pig, Hive, Hbase, Mahouth. SQL Essentials Joins, Set Operations, Grouping Extensions. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 10 | | | | | | |
| **4** | | | | HADOOP DISTRIBUTED FILE SYSTEM ARCHITECTURE: HDFS Architecture, HDFS Concepts, Blocks, NameNode, Secondary NameNode, DataNode, HDFS Federation, HDFS High Availability, Basic File System Operations, Data Flow, Anatomy of File Read, Anatomy of File Write, Anatomy of a MapReduce Job Run . | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 10 | | | | | | |
| **5** | | | | PROCESSING YOUR DATA WITH MAPREDUCE: Getting to know MapReduce, MapReduce Execution Pipeline, Runtime Coordination and Task Management, MapReduce Application, Hadoop Word Count Implementation, Installing and Running Pig, Hbase Versus RDBMS, Installing and Running ZooKeeper. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 10 | | | | | | |
|  | | | | **TOTAL** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | **48** | | | | | | |
|  | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  | | | | | | |
|  | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  | | | | | | |
| **11** | | **Suggested Books:** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  | | | | | | |
| **Sl. NO.** | | | **NAME OF AUTHORS/BOOKS/PUBLISHERS** | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | **YEAR OF PUBLICATION** | | | | | |
| **1** | | | David Dietrich, Barry Heller and Beibei Yang, “Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data”, EMC Education Services, Reprint 2015, Wiley, ISBN: 9788126556533. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2015 | | | | | |
| **2** | | | Tom White, "Hadoop: The Definitive Guide", 4th Edition, 2015, O'Reilly, ISBN: 9789352130672. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2015 | | | | | |
| **3** | | | BirisLublinsky, Kevin T. Smith and Alexey Yakubovich, “Professional Hadoop Solutions”, Reprint 2014, Wiley, ISBN 13:9788126551071. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2014 | | | | | |
| **4** | | | Stephen Marsland, “Machine Learning – An Algorithmic Perspective”, , Taylor& Francis Group, Second Edition, 2015, Chapman & Hall / CRC Press , ISBN:9781466583283. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2015 | | | | | |
| **5** | | | Nathan Marz, James Warren, “Big Data-Principles and best practices of scalable real-time data systems”, Edition 2015, DreamTech Press, ISBN: 9789351198062. | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | 2015 | | | | | |