

<b>Sub. Code</b>	<b>Subject Name</b>	<b>Credit</b>	<b>Theory</b>	<b>Laboratory</b>
IT-VAC-501	Learning Big Data Analysis Through WEKA	2	50 [Continuous Evaluation]	50

<b>Objectives</b>	<p>The objective of this course is to learn the basic concepts of Big Data including various methods for analysis of data with 3Vs.</p> <p>This course is designed for students who have no previous knowledge of dataanalysis but wish to acquire these skills in a short period of time. These studentswill learn how to analyze large data sets and identify patterns that will improveany company’s and organization decision-making process.</p>
<b>Pre-Requisites</b>	<p>Interested students should have prior knowledge on</p> <ul style="list-style-type: none"> <li>• Science (testing and formulating hypothesis)</li> <li>• Statistical methodology</li> <li>• Programming</li> <li>• Data mining</li> </ul>
<b>Course Outcome</b>	<p>At the end of the course, the students will be able to:</p> <ol style="list-style-type: none"> <li>i) Capture, categorize, simplify, normalize and prepare data to be processed</li> <li>ii) Work with and analyze large data sets</li> <li>iii) Visually represent analysis’s conclusions to technical and non technicalaudiences</li> <li>iv) Use the most common algorithms, to make sense of large amounts of data,which are applicable to most business and management problems.</li> <li>v) Learn WEKA and apply machine learning techniques using R.</li> </ol>
<b>Teaching Scheme</b>	<p>Regular classroom lectures with use of ICT tools as and when required along with Laboratory classes for enhancing domain specific skill.</p>

### **Detailed Syllabus**

<b>Unit</b>	<b>Topics</b>	<b>Hours</b>
-------------	---------------	--------------

I	Types of Digital Data, Introduction to Big Data, Big Data Analytics, History of WEKA, Analysing Data with WEKA.	6
II	Introduction to Big data tools like Hadoop, Map reduce, and HDFS	6
III	Large-scale graph processing techniques, Big data stream techniques and algorithms, Large-scale probabilistic data analysis	9
IV	Machine Learning: Introduction, Supervised Learning, Unsupervised Learning	9
V	Recommendation system, Big Data Analysis through R	
Total		30

**Text Books:**

T1.R. Buyya, R. Calheiros, and A.V. Dastjerdi, “Big Data: Principles and Paradigms”, Morgan Kaufmann, 2016.

Web reference: WEKA for Big Data analysis  
(<https://www.cs.waikato.ac.nz/ml/weka/index.html>)

**Reference Books:**

R1. B. S.P. Mishra, Satchidananda Dehuri, et al. (Eds.), Cloud Computing for Optimization: Foundations, Applications, and Challenges, Springer, 2018.

R2.D.P. Acharjya , Satchidananda Dehuri, et al. (Eds.), Computational Intelligence for Big Data Analysis: Frontier Advances and Applications, Springer, 2016