# Manipal School of Information Sciences (MSIS) \*\*Manipal Academy of Higher Education, Manipal Master of Engineering - ME (Big Data Analytics)\*\*## \*\*Course File\*\* Course Name : Algorithm and Data Structures for Big Data Lab

--- --- ---

Course Code : BDA 5151

Academic Year : 2024 – 2025

Semester : I

Name of the Course Coordinator : Mr. DEEPAK RAO B

Name of the Program Coordinator : Dr. PRATHVIRAJ N Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date 1. Course Plan 5

--- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs) 5

1.3 Course Content (Syllabus) Error! Bookmark not defined.

1.4 References Error! Bookmark not defined.

1.5 Other Resources (Online, Text, Multimedia, etc.) 9

1.6 Topic Learning Outcomes (TLOs) Error! Bookmark not defined.

1.7 Course Timetable 9

1.8 Assessment Plan 9

1.9 Assessment Details 11

1.10 Mapping of COs with POs Error! Bookmark not defined.

2. Assessment Details Error! Bookmark not defined.

2.1 Student Details: Error! Bookmark not defined.

2.2 Assessment outcomes Error! Bookmark not defined.

2.3 Analysis of Assessment outcomes Error! Bookmark not defined.

2.4 Attainment of Course Outcomes (Direct) Error! Bookmark not defined. 2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire Error! Bookmark not defined.

--- --- --- ---

2.6 Attainment of Course Outcomes (Indirect): Analysis Error! Bookmark not defined.

3. CO-PO Assessment Error! Bookmark not defined.

4. Observations and Comments Error! Bookmark not defined.

4.1 Observations from Course Coordinator based on the direct assessment Error! Bookmark not defined.

4.2 Comments/Suggestions by the Course Coordinator Error! Bookmark not defined. # Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences. # Program Outcomes (POs)By the end of the postgraduate program in Big Data Analytics, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program. PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

--- ---

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. # 1. Course Plan### 1.1 Primary Information Course Name : Algorithms and Data Structures for Big Data Lab

--- --- ---

L-T-P-C : 0-0-3-1

Contact Hours : 36 Hours

Pre-requisite : Programming with Python or C #### 1.2 Course Outcomes (COs) CO At the end of this course, the student should be able No. of Contact Program Outcomes BL

--- --- --- --- ---

to: Hours (PO's)

Design programs for implementation of linked lists, stack

CO1 and queues. 15 PO4 5 Design programs for implementation of binary search

--- --- --- --- ---

CO2 tree, sorting and searching, dictionary and Hash Table 12 PO4 5

Design programs for graphs and shortest path techniques.

CO3 9 PO4 5 #### 1.3 Assessment Plan Components Lab Test Flexible Assessments End semester/ Makeup

--- --- --- ---

(5- 6 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of Applying; Analyzing. Applying; Analyzing. Applying; Analyzing; Evaluating.

questions Evaluating. Pattern Answer all the Assignment: develop applications using Answer all the questions. Maximum

--- --- --- ---

questions. Maximum various data structures and different marks 50.

marks 30. design techniques

Schedule As per academic Assignment submission: November As per academic calendar.

calendar. 2024

Topics covered Linked List, Stack,

Queue, Trees, Searching Comprehensive examination

& Sorting, Hash tables, covering the full syllabus.

Graphs #### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO ---

mapping, reference books

Lab1 Linked List: Implementing Single Linked List CO1

Lab2 Linked List: Implementing Double Linked List CO1

Lab3 Linked List: Application development using linked lists CO1 Lab4 Stack: Implementation and applications of Stack CO1

--- --- ---

Lab5 Queue: Implementation and applications of Queue CO1

Lab6 Tree: Implementation and applications of Tree CO2

IT1 Internal lab test CO1, CO2

Lab7 Applications using different search and sorting techniques. CO2

Lab8 Applications using different search and sorting techniques. CO2

Lab9 Application using Hash Table CO2

Lab10 Graph representation using list and matrix method CO3

Lab11 Graph traversal CO3

Lab12 Graph: Shortest path technique CO3 #### 1.5 References- 1. Introduction to Algorithms Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press.

- 2. Data Structures and Algorithms Aho, Hopcroft and Ulmann. Pearson Publishers.

- 3. Data Structures and Algorithms in Python Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons.

- 4. Data Streams: Algorithms and Applications S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236## 1.6 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.#### 1.7 Course Timetable st 1 Semester Big Data Analytics Room: LG1 LH 3 Lab: Data Science Lab

--- --- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE

WED

THU

FRI ADS LAB

SAT 1.8 Assessment Plan COs Marks & weightage

--- --- --- --- --- ---

CO No. CO Name Lab Test Assignment End Semester CO wise

(Max. 30) (Max. 20) (Max. 50) Weightage

CO1 Design programs for implementation of linked lists, stack and queues. 15 10 25 0.5

CO2 Design programs for implementation of binary search tree, sorting and searching, dictionary and Hash Table 10 6 15 0.31

CO3 Design programs for graphs and shortest path techniques. 5 4 10 0.19

Marks (weightage) 0.3 0.2 0.5 1.0 - In-semester Assessment is considered as the Internal Assessment (IA) in each subject for 50 marks, which includes the performances in class / tutorial participation, assignment work, lab work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for each lab subject is conducted for a maximum of 50.

- End-semester mark for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in a subject.### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. Tools (TLP) Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- --- ---

No.

1 Sessional 0.3 2 • • Performance is measured using sessional attainment level. Reference: question paper and answer scheme.

• Each sessional is assessed for a maximum of 30 marks.

2 Assignments 0.2 - • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 20 marks.

3 ESE 0.5 1 • • Performance is measured using ESE attainment level. Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 50 mark. # 1.10 Course Articulation Matrix CO1 Y

--- --- --- --- --- ---

CO CO2 PO1 PO2 PO3 PO4 Y PO5 CO3 Y

--- --- --- ---

Average Articulation Level \* \* \* ![](\_page\_0\_Picture\_0.jpeg)# Master of Engineering - ME (Big Data Analytics) Course Name : Principles of Data Visualization

--- --- ---

Course Code : BDA 5132

Academic Year : 2024 - 25

Semester : I

Name of the Course Coordinator : SATYANARAYAN SHENOY

Name of the Program Coordinator : Dr. PRATHVIRAJ N #### Course File Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date Table of Contents

--- ---

1. Course Plan 6

1.1 Primary Information 6

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping 7

1.3 Assessment Plan 8

1.4 Lesson Plan 9

1.5 References 11

1.6 Other Resources (Online, Text, Multimedia, etc.) 11

1.7 Course Timetable 12

1.8 Assessment Plan 13

1.9 Assessment Details 15

1.10 Course Articulation Matrix 16

2. Assessment Details 17

2.1 Student Details: 17

2.2 Assessment outcomes 19

2.3 Analysis of Assessment outcomes 21 2.4 Attainment of Course Outcomes (Direct) 22

--- ---

2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire 23

2.6 Attainment of Course Outcomes (Indirect): Analysis 24

3. CO-PO Assessment 25

4. Observations and Comments 26

4.1 Observations from Course Coordinator based on the direct and indirect assessments 26

4.2 Comments/Suggestions by the Course Coordinator 27 ## Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for ME (Big Data Analytics), program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and

applied research experiences. ## Program Outcomes (POs)By the end of the postgraduate program in ME (Big Data Analytics), graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. ![](\_page\_5\_Picture\_0.jpeg)## 1. Course Plan### 1.1 Primary Information Course Name : Principles of Data Visualization [BDA 5132]

--- --- ---

L-T-P-C : 3-0-0-3

Contact Hours : 36 Hours

Pre-requisite : Basic Programming with Python

Core/ PE/OE : Elective ![](\_page\_6\_Picture\_0.jpeg)### 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

Implement web scrapping techniques to extract data from

CO1 websites. 12 PO4 3

CO2 Organize raw data for analysis using data manipulation techniques. 12 PO1 4

CO3 Use powerBI for preparation and modelling of data for analysis. 4 PO4 3

CO4 Interpret data using various data visualization techniques. 4 PO5 3

Report data for analytics and to manage workspace using

CO5 Power BI 4 PO3 2 ![](\_page\_7\_Picture\_0.jpeg)#### 1.3 Assessment Plan Components Internal Test 1 Flexible Assessments (2 – 3 in number) End semester/ Makeup examination

--- --- --- ---

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of questions Applying; Analyzing. Applying; Analyzing. Understanding. Applying; Analyzing; Understanding.

Pattern Answer all 5 questions of 10 marks each. Assignment: (Solving Use case using scraping and visualization techniques.) Answer all 10 full questions of 10 marks each.

Schedule As per academic calendar. Assignment submission: November 2024 As per academic calendar.

Topics covered Web scraping, Data Analysis, Data Analysis, Data Web scraping, Data Analysis and Data Visualization. Comprehensive examination covering the full syllabus. Students are expected to

Visualization answer all questions. ![](\_page\_8\_Picture\_0.jpeg)#### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

L1 Web scraping: Introduction CO1

L2 Web scraping models and techniques CO1

L3 Web scraping models and techniques CO1

L4 Usecases: BeautifulSoup CO1

L5 Usecases: Scrapy CO1

L6 Usecases: Scrapy CO1

L7 Data Analysis: introduction CO2

L8 Data structures for analysis: numpy CO2

L9 Data structures for analysis: numpy CO2

L10 Data Structures for analysis: Pandas CO2

L11 Data Structures for analysis: Pandas CO2

L12 Data Wrangling - Clean, Transform, Merge, Reshape CO2

L13 Data Wrangling - Clean, Transform, Merge, Reshape CO2 ![](\_page\_9\_Picture\_0.jpeg) L14 Data Aggregation and Group Operations CO2

--- --- ---

L15 Power BI Introduction CO3

L16 Prepare data for analysis using Power BI CO3

L17 Prepare data for analysis using Power BI CO3

L18 Model data in Power BI CO3

IT1 Internal test 1 CO1, CO2, CO3

L19 Data Visualization: Introduction CO4

L20 Visualization techniques: time series CO4

L21 Visualization techniques: time series CO4

L22 Visualization techniques: statistical distributions CO4

L23 Visualization techniques: statistical distributions CO4

L24 Visualization techniques: statistical distributions CO4

L25 Visualization techniques: maps - Data visualization for web. CO4

L26 Visualization techniques: maps - Data visualization for web. CO4

L27 Visualization techniques: maps - Data visualization for web. CO4

L28 Visualize data in Power BI. CO5

L29 Data analysis in Power BI CO5

L30 Data analysis in Power BI CO5

L31 Manage workspaces and datasets in Power BI CO5

L32 Manage workspaces and datasets in Power BI CO5 ![](\_page\_10\_Picture\_0.jpeg) L33 Create and use analytics reports with Power BI CO5

--- --- ---

L34 Create and use analytics reports with Power BI CO5

L35 Manage workspaces power BI CO5

L36 Manage workspaces power BI CO5 ### 1.5 References- 1. Website Scraping with Python: Using BeautifulSoup and Scrapy, Gábor & Hajba, APRESS Publications, 1st Edition, 2018.

- 2. Web Scraping with Python: Collecting More Data from the Modern Web, Ryan Mitchell Shroff, O'Reilly, 2nd Edition, 2018.

- 3. Designing Data Visualizations, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011.

- 4. Python for Data Analysis, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.

- 5. https://learn.microsoft.com/en-us/certifications/exams/pl-300/### 5.1 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.![](\_page\_11\_Picture\_0.jpeg)### 5.2 Course Timetable 1 st Semester Big Data Analytics Room: LG1 LH 8

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE PDV

WED

THU PDV LAB PDV

FRI

SAT PDV ![](\_page\_12\_Picture\_0.jpeg)#### 5.3 Assessment Plan Cos Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name Mid Semester Assignment End Semester CO wise

(Max. 50) (Max. 20) (Max. 100) Weightage

CO1 Implement web scrapping techniques to extract data from websites. 20 2 20 0.2

CO2 Organize raw data for analysis using data 30 2 30 0.29

manipulation techniques.

CO3 Use powerBI for preparation and modelling of data - 2 10 0.15

for analysis.

CO4 Interpret data using various data visualization - 2 30 .25

techniques.

CO5 Report data for analytics and to manage workspace - 2 10 .11

using Power BI

Marks (weightage) 0.2 0.1 0.5 1.0 Note:- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.

- End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.Weightage for CO1 = (IT1 marks for CO1 / 2.5 + IT2 marks for CO1 / 2.5 + Assignment marks for CO1 + ESE marks for CO1 / 2)/100 = (25/2.5 + 0 + 0 + 20/ 2)/100 = 0.2![](\_page\_14\_Picture\_0.jpeg)#### 5.4 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

 Performance is measured using internal test attainment level.

 Reference: question paper and answer scheme.

1 Internal Test 0.4 2  Each internal test is assessed for a maximum of 50 marks and scaled down to 40

marks.

2 Assignments 0.1  1 Performance is measured using assignments/quiz attainment level.

 Assignments/quiz are evaluated for a maximum of 10 marks.

 Performance is measured using ESE attainment level.

3 ESE 0.5 1  Reference: question paper and answer scheme.

 ESE is assessed for a maximum of 100 marks and scaled down to 50 marks. ![](\_page\_15\_Picture\_0.jpeg)### 5.5 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y Y

CO3 Y

CO4 Y

CO5 Y

Average Articulation Level \* \* \* \* # \*\*Master of Engineering - ME (Big Data Analytics)\*\* Course Name : Fundamentals of Machine Learning Lab

--- --- ---

Course Code : BDA 5153

Academic Year : 2024 - 25

Semester : I

Name of the Course Coordinator : Dr. Arockiaraj S

Name of the Program Coordinator : Dr. Prathviraj N #### \*\*Course File\*\* Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date 1. Course Plan 5

--- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping 6

1.3 Assessment Plan 7

1.4 Lesson Plan 8

1.5 References 9

1.6 Other Resources (Online, Text, Multimedia, etc.) 9

1.7 Course Timetable 10

1.8 Assessment Plan 11

1.9 Assessment Details 12

1.10 Course Articulation Matrix 13 ## Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences. ## Program Outcomes (POs)By the end of the postgraduate program in \*\*ME (Big Data Analytics)\*\*, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. ## 1. Course Plan#### 1.1 Primary Information Course Name : Fundamentals of Machine Learning Lab [BDA 5153]

--- --- ---

L-T-P-C : 0-0-3-1

Contact Hours : 36 Hours

Pre-requisite : Basic Programming with Python

Core/ PE/OE : Core ### 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

CO1 Apply different types of supervised and unsupervised 8 PO3 3

machine learning algorithms to practical problems.

CO2 Analyse different types of machine learning paradigms. 12 PO4 4

CO3 Evaluate the performance of machine learning algorithms. 16 PO5 5 #### 1.3 Assessment Plan Components Lab test Flexible Assessments End semester/ Makeup

--- --- --- --- ---

(2 – 3 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of Applying; Analyzing and Evaluating. Applying; Analyzing. Applying; Analyzing;

questions Evaluating. Evaluating.

Assignment: Solving problems by

Pattern Answer all the questions. Maximum marks applying, analyzing and evaluating Answer all the questions.

30. Generative AI use cases. Maximum marks 50.

[To be decided by the faculty.]

Schedule To be decided by the faculty Assignment submission: (To be decided by the faculty) To be decided by the faculty

Introduction to Machine Learning; Decision Comprehensive examination

Topics Trees- Linear Model: K-nearest Neighbours Lab assignment: Implement pre-training, covering the full syllabus.

fine-tuning, and evaluation of large

covered Algorithm- Cross-validation - Dimension language models. Students are expected to

Reduction. answer all questions. #### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

L1 Program data ingestion, perform data wrangling, understand the data matrix, and differentiate between sample and feature. CO1

L2 Implement decision tree models in Python, fine-tune model parameters, and interpret results. CO1

L3 Implement linear models in Python and interpret model coefficients for practical problems. CO2

L4 Implement and visualize bias-variance trade-off using linear regression as a basis; Implement the K-nearest neighbours algorithm. CO2

L5 Implement, visualize, compare, contrast, and interpret the results of dimension reduction applied to practical data using PCA, MDS, and t-SNE. CO2

IT1 Internal lab test CO1 & CO2

L6 Through coding, understand how ensemble methods in machine learning work. CO3

L7 Implement maximum likelihood estimation for a simple model. CO3

L8 Analyse the performance of the Naive Bayes and logistic regression models for practical problems using appropriate performance metrics. CO3 L9 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

--- --- ---

L10 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

L11 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3

L12 Lab assignment: Implement pre-training, fine-tuning, and evaluation of large language models. CO3 #### 1.5 References- 1. Module: Introduction to Machine Learning (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academicprogram/educators/intro-machine-learning-training-kit.html)

- 2. Module: Get started with AI on Azure (https://learn.microsoft.com/en-us/training/modules/get-started-ai-fundamentals/)

- 3. Module: Microsoft Azure AI Fundamentals: Get started with artificial intelligence (https://learn.microsoft.com/en-us/training/paths/getstarted-with-artificial-intelligence-on-azure/)

- 4. Learning path: Understand data science for machine learning (https://learn.microsoft.com/en-us/training/paths/understand-machinelearning/)

- 5. Module: Generative AI with Large Language Models (https://www.coursera.org/learn/generative-ai-with-llms)

- 6. Grokking Machine Learning, Luis G. Serrano, Manning Publications; 1st Edition, 2019 Online resource from Manning Publications available at https://www.manning.com/books/grokking-machine-learning

- 7. A Course in Machine Learning, Hal Daumé III Online resource available at http://ciml.info/

- 8. An Introduction to Statistical Learning: with Applications in Python (Springer Texts in Statistics), Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor, 1st Edition, 2023 – Online resource available at https://www.statlearning.com/### 1.6 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.#### 1.7 Course Timetable 1 st Semester Big Data Analytics Lab:

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE FML LAB

WED

THU

FRI

SAT #### 1.8 Assessment Plan COs Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name Mid semester Assignment End Semester CO wise

(Max. 50) (Max. 20) (Max. 100) Weightage

CO1 Apply different types of supervised and unsupervised 25 5 20 0.30

machine learning algorithms to practical problems.

Analyse different types of machine learning

CO2 25 5 40 0.40

paradigms.

Evaluate the performance of machine learning

CO3 - 10 40 0.30

algorithms.

Marks (weightage) 0.3 0.2 0.5 1.0 Note:- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.

- End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.Weightage for CO1 = (mid semester marks for CO1 / 1.6666 + Assignment marks for CO1/1.0 + ESE marks for CO1 / 2)/100#### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

• Performance is measured using internal test attainment level.

1 Internal Test 0.3 1 • Reference: question paper and answer scheme.

• Each internal test is assessed for a maximum of 50 marks and scaled down to 40

marks.

2 Assignments 0.2 2 • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 10 marks.

• Performance is measured using ESE attainment level.

3 ESE 0.5 1 • Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 100 marks and scaled down to 50 marks. #### 1.10 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y

CO3 Y

Average Articulation Level Y Y Y # Manipal School of Information Sciences (MSIS) \*\*Manipal Academy of Higher Education, Manipal Master of Engineering - ME (Big Data Analytics)\*\*## \*\*Course File\*\* Course Name : Algorithms and Data Structures for Big Data

--- --- ---

Course Code : BDA 5101

Academic Year : 2024 – 2025

Semester : I

Name of the Course Coordinator : Mr. DEEPAK RAO B

Name of the Program Coordinator : Dr. PRATHVIRAJ N Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date 1. Course Plan 5

--- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs) Error! Bookmark not defined.

1.3 Course Content (Syllabus) Error! Bookmark not defined.

1.4 References Error! Bookmark not defined.

1.5 Other Resources (Online, Text, Multimedia, etc.) 10

1.6 Topic Learning Outcomes (TLOs) Error! Bookmark not defined.

1.7 Course Timetable 11

1.8 Assessment Plan 11

1.9 Assessment Details 12

1.10 Mapping of COs with POs Error! Bookmark not defined.

2. Assessment Details Error! Bookmark not defined.

2.1 Student Details: Error! Bookmark not defined.

2.2 Assessment outcomes Error! Bookmark not defined.

2.3 Analysis of Assessment outcomes Error! Bookmark not defined.

2.4 Attainment of Course Outcomes (Direct) Error! Bookmark not defined.

2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire Error! Bookmark not defined.

2.6 Attainment of Course Outcomes (Indirect): Analysis Error! Bookmark not defined. - \*\*3. CO-PO Assessment.........................................\*\*Error! Bookmark not defined.

- \*\*4. Observations and Comments.........................\*\*Error! Bookmark not defined.

- \*\*4.1 Observations from Course Coordinator based on the direct assessment ................................................\*\* Error! Bookmark not defined.

- \*\*4.2 Comments/Suggestions by the Course Coordinator ..................................................................................\*\* Error! Bookmark not defined.# Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied

research experiences. # Program Outcomes (POs)By the end of the postgraduate program in Big Data Analytics, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data. # 1. Course Plan## 1.1 Primary Information Course Name : Algorithms and Data Structures for Big Data

--- --- ---

L-T-P-C : 3-0-0-3

Contact Hours : 36 Hours

Pre-requisite : Programming with Python or C

Core/ PE/OE : Core # 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be No. of Contact Marks Program Outcomes BL

--- --- --- --- --- ---

able to: Hours (PO's)

CO1 Analyze recursive programs, solve a general class of recurrence relations 3 10 PO3 3

CO2 Design programs for implementation of linked lists, 10 40 PO4 4

stack, queues and binary search tree CO3 Design programs for dictionary, hash tables, graphs, 14 40 PO4 4

--- --- --- --- --- ---

shortest path techniques, sorting and searching.

CO4 Design string and text processing programs. 9 10 PO4 4

Total 100 #### 1.3 Assessment Plan Components Mid Semester Flexible Assessments End semester/ Makeup

--- --- --- ---

(2 – 3 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 30% 20% 50%

Typology of Applying; Analyzing and Applying; Analyzing. Applying; Analyzing;

questions Evaluating Evaluating. Evaluating. Answer all 5 Quiz / Test: 10%

--- --- --- ---

questions of 10 Assignment 1: (Text processing Answer all 10 full questions of

marks each. Each algorithms, 5% weightage) 10 marks each. Each question

Pattern question may have Assignment 2: (Data Streaming may have 2 to 3 parts of

2 to 3 parts of algorithms, 5% weightage) 3/4/5/6/7 marks.

3/4/5/6/7 marks.

As per academic Quiz / Test: September 2024

Schedule calendar. Assignment 1: October 2024 As per academic calendar.

Assignment 2: November 2024

Topics covered Algorithm Analysis Techniques Elementary Data structures

Comprehensive examination

covering the full syllabus.

Trees Students are expected to

Sorting, Searching

answer all questions.

Dictionary and

Hashing #### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO ---

mapping, reference books

L1 Analysis of recursive programs. CO1

L2 Solving recurrence equations. CO1

L3 General solution for a large class of recurrences. CO1

L4 Implementation of lists. CO2

L5 Implementation of lists. CO2

L6 Implementation of lists. CO2

L7 Implementation of stacks. CO2

L8 Implementation of stacks. CO2

L9 Implementation of queues. CO2

L10 Implementation of queues. CO2

L11 Implementation of Trees. CO2

L12 Implementation of Trees. CO2

L13 Implementation of Trees. CO2

L14 Sorting – bubble, selection and insertion CO3

L15 Sorting – Quick sort CO3

L16 Sorting – Merge sort CO3

L17 Sorting – Heap Sort CO3 L18 Searching – Linear and binary CO3

--- --- ---

L19 Dictionary and Hash Tables CO3

L20 Dictionary and Hash Tables CO3

L21 Dictionary and Hash Tables CO3

Mid Semester Evaluation CO1 , CO2, CO3

L22 Graph Terminology CO3

L23 Representation of graphs CO3

L24 Traversing Graphs CO3

L25 Shortest Path algorithm CO3

L26 Shortest Path algorithm CO3

L27 String and Text processing CO4

L28 String and Text processing CO4

L29 String and Text processing CO4

L30 String and Text processing CO4

L31 Data streaming algorithms CO4

L32 Data streaming algorithms CO4

L33 Data streaming algorithms CO4

L34 Data streaming algorithms CO4

L35 Data streaming algorithms CO4

L36 Data streaming algorithms CO4 #### 1.5 References- 1. Introduction to Algorithms Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press.

- 2. Data Structures and Algorithms Aho, Hopcroft and Ulmann. Pearson Publishers.

- 3. Data Structures and Algorithms in Python Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons.

- 4. Data Streams: Algorithms and Applications S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236

- 5. https://in.coursera.org/specializations/boulder-data-structures-algorithms#### 1.6 Other Resources (Online, Text, Multimedia, etc.)- 6. Web Resources: Blog, Online tools and cloud resources.

- 7. Journal Articles.### 1.7 Course Timetable st 1 Semester Big Data Analytics Room: LG1 LH 8 Lab: Data Science Lab

--- --- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON ADS

TUE

WED ADS

THU

FRI ADS

SAT #### 1.8 Assessment Plan Cos Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name IT-1 Assignment End Semester CO wise

(Max. 50) (Max. 10) (Max. 100) Weightage

CO1 Analyze recursive programs, solve a general class of 5 - 10 0.09

recurrence relations

CO2 Design programs for implementation of linked lists, 30 10 50 0.53

stack, queues, binary search tree Design for and searching, programs sorting

--- --- --- --- --- --- --- ---

CO3 dictionary, hash tables, graphs and shortest path 15 40 0.32

techniques.

CO4 Design string and text processing programs. - 10 0.06

Marks (weightage) 0.3 0.2 0.5 1.0 Note:- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.#### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools (TLP) Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

1 Mid Semester 0.3 • 1 Performance is measured using sessional attainment level.

• Reference: question paper and answer scheme. • Mid semester exam is assessed for a maximum of 50 marks and scaled down to

--- --- --- --- --- ---

30 marks

2 Assignments 0.2 1 • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 20 marks.

• Performance is measured using ESE attainment level.

3 ESE 0.5 1 • Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 100 marks and scaled down to 50 marks. ### 1.10 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y

CO3 Y

CO4 Y

Average Articulation Level \* \* ![](\_page\_0\_Picture\_0.jpeg)# Master of Engineering - ME (Big Data Analytics) Course Name : Principles of Data Visualization Lab

--- --- ---

Course Code : BDA 5182

Academic Year : 2024 - 25

Semester : I

Name of the Course Coordinator : SATYANARAYAN SHENOY

Name of the Program Coordinator : Dr. PRATHVIRAJ N ### Course File Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date Table of Contents

--- ---

1. Course Plan 6

1.1 Primary Information 6

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping 7

1.3 Assessment Plan 8

1.4 Lesson Plan 9

1.5 References 10

1.6 Other Resources (Online, Text, Multimedia, etc.) 10

1.7 Course Timetable 11

1.8 Assessment Plan 12

1.9 Assessment Details 14

1.10 Course Articulation Matrix 15

2. Assessment Details 16

2.1 Student Details: 16

2.2 Assessment outcomes 18

2.3 Analysis of Assessment outcomes 20 2.4 Attainment of Course Outcomes (Direct) 21

--- ---

2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire 22

2.6 Attainment of Course Outcomes (Indirect): Analysis 23

3. CO-PO Assessment 24

4. Observations and Comments 25

4.1 Observations from Course Coordinator based on the direct and indirect assessments 25

4.2 Comments/Suggestions by the Course Coordinator 26 # Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for ME (Big Data Analytics), program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and

applied research experiences. # Program Outcomes (POs)By the end of the postgraduate program in ME (Big Data Analytics), graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. ![](\_page\_5\_Picture\_0.jpeg)# 1. Course Plan# 1.1 Primary Information Course Name : Principles of Data Visualization Lab [BDA 5132]

--- --- ---

L-T-P-C : 0-0-3-1

Contact Hours : 36 Hours

Pre-requisite : Programming with Python, Power BI

Core/ PE/OE : Elective ![](\_page\_6\_Picture\_0.jpeg)# 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

Experiment web scrapping techniques to extract data from

CO1 websites. 9 PO4 4

CO2 Implement NumPy and Pandas for data science operations with examples. 6 PO3 3

CO3 Organize data for visualization using data manipulation techniques. 6 PO4 4

CO4 Experiment different visualization techniques 6 PO4 4

CO5 Use power BI for analytics and to manage workspace. 9 PO5 3 ![](\_page\_7\_Picture\_0.jpeg)## 1.3 Assessment Plan Components Lab Test Flexible Assessments (2 – 3 in number) End semester/ Makeup examination

--- --- --- ---

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of questions Applying; Analyzing. Applying; Analyzing. Applying; Analyzing.

Pattern Answer all the questions. Maximum marks 30. Assignment: (Solving Use case using scraping and visualization techniques.) Answer all the questions. Maximum marks 50

Schedule As per academic calendar. Assignment submission: November 2024 As per academic calendar.

Topics covered Scraping tools, NumPy, Pandas, PowerBI Comprehensive examination covering the full syllabus. ![](\_page\_8\_Picture\_0.jpeg)## 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

Lab 1 Web scraping using Beautiful soups CO1

Lab 2 Web scraping using Scrapy framework CO1

Lab 3 Web scraping using Scrapy framework CO1

Lab 4 Panda, NumPy: Implement NumPy and Pandas for data science operations with examples. CO2

Lab 5 Panda, NumPy: Implement NumPy and Pandas for data science operations with examples. CO2

Lab 6 Data Wrangling - Clean, Transform, Merge, Reshape CO3

IT1 IT1 Internal lab test CO1, CO2, CO3

Lab 7 Data Aggregation and Group Operations. CO3

Lab 8 Visualization techniques: time series, statistical distributions. CO4

Lab 9 Visualization techniques: maps - Data visualization for web. CO4

Lab 10 Visualize data and analysis in Power BI CO5

Lab 11 Manage workspaces and datasets in Power BI. CO5

Lab 12 Create and use analytics reports with Power BI. CO5 # 1.5 References- 1. Website Scraping with Python: Using BeautifulSoup and Scrapy, Gábor & Hajba, APRESS Publications, 1st Edition, 2018.

- 2. Web Scraping with Python: Collecting More Data from the Modern Web, Ryan Mitchell Shroff, O'Reilly, 2nd Edition, 2018.

- 3. Designing Data Visualizations, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011.

- 4. Python for Data Analysis, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.

- 5. https://learn.microsoft.com/en-us/certifications/exams/pl-300/## 1.6 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.![](\_page\_10\_Picture\_0.jpeg)# 1.7 Course Timetable 1 st Semester Big Data Analytics Room: LG1 LH 8 Lab: Data Science Lab

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE PDV

WED

THU PDV LAB PDV

FRI

SAT PDV ![](\_page\_11\_Picture\_0.jpeg)### 1.8 Assessment Plan Cos Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name Mid Semester Assignment End Semester CO wise

(Max. 50) (Max. 20) (Max. 100) Weightage

CO1 Experiment web scrapping techniques to extract 10 4 20 0.34

data from websites.

CO2 Implement NumPy and Pandas for data science 10 4 - 0.14

operations with examples.

CO3 Organize data for visualization using data - 4 20 0.24

manipulation techniques.

CO4 Experiment different visualization techniques. 10 4 - 0.14

CO5 Use power BI for analytics and to manage - 4 10 0.14

workspace.

Marks (weightage) 0.3 0.2 0.5 1.0 Note:- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in lab participation, assignment work, lab work, lab tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 50.

- End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.Weightage for CO1 = (Lab Test marks for CO1 + Assignment marks for CO1 + ESE marks for CO1) /100 = (5 + 2 +5)/100 = 0.12![](\_page\_13\_Picture\_0.jpeg)### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

 Performance is measured using lab internal test attainment level.

1 Internal Test 0.3 1  Reference: question paper and answer scheme.

 Lab internal test is assessed for a maximum of 30 marks.

2 Assignments 0.2  1 Performance is measured using assignments attainment level.

 Assignment is evaluated for a maximum of 20 marks.

 Performance is measured using ESE attainment level.

3 ESE 0.5 1  Reference: question paper and answer scheme.

 ESE is assessed for a maximum of 50 marks. ![](\_page\_14\_Picture\_0.jpeg)# 1.10 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y

CO3 Y

CO4 Y

CO5 Y

Average Articulation Level \* \* \* # Manipal School of Information Sciences (MSIS) \*\*Manipal Academy of Higher Education, Manipal Master of Engineering - ME (Big Data Analytics)\*\*## \*\*Course File\*\* Course Name : Architecture of Big Data Systems Lab

--- --- ---

Course Code : BDA 5152

Academic Year : 2024 – 2025

Semester : I

Name of the Course Coordinator : Mr. DEEPAK RAO B

Name of the Program Coordinator : Dr. PRATHVIRAJ N Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date 1. Course Plan 5

--- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs) 5

1.3 Course Content (Syllabus) Error! Bookmark not defined.

1.4 References Error! Bookmark not defined.

1.5 Other Resources (Online, Text, Multimedia, etc.) 9

1.6 Topic Learning Outcomes (TLOs) Error! Bookmark not defined.

1.7 Course Timetable Error! Bookmark not defined.

1.8 Assessment Plan 10

1.9 Assessment Details 11

1.10 Mapping of COs with POs Error! Bookmark not defined.

2. Assessment Details Error! Bookmark not defined.

2.1 Student Details: Error! Bookmark not defined.

2.2 Assessment outcomes Error! Bookmark not defined.

2.3 Analysis of Assessment outcomes Error! Bookmark not defined.

2.4 Attainment of Course Outcomes (Direct) Error! Bookmark not defined. 2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire Error! Bookmark not defined.

--- --- --- ---

2.6 Attainment of Course Outcomes (Indirect): Analysis Error! Bookmark not defined.

3. CO-PO Assessment Error! Bookmark not defined.

4. Observations and Comments Error! Bookmark not defined.

4.1 Observations from Course Coordinator based on the direct assessment Error! Bookmark not defined.

4.2 Comments/Suggestions by the Course Coordinator Error! Bookmark not defined. # Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences. # Program Outcomes (POs)By the end of the postgraduate program in Big Data Analytics, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program. PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

--- ---

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. # 1. Course Plan### 1.1 Primary Information Course Name : Architecture of Bid Data Systems Lab

--- --- ---

L-T-P-C : 0-0-3-1

Contact Hours : 36 Hours

Pre-requisite : Programming with Python or Java ### 1.2 Course Outcomes (COs) CO At the end of this course, the student should be able No. of Contact Program Outcomes BL

--- --- --- --- ---

to: Hours (PO's) Use data extraction tools to ingest various types of data

--- --- --- --- ---

CO1 into big data systems. 6 PO3 3

Experiment with different tools and frameworks of

CO2 Hadoop eco-system. 12 PO4 4

CO3 Experiment with Spark Engine to process real-time data. 12 PO4 4

Design applications to handle batch and streaming data

CO4 using Hadoop and Spark tools. 6 PO5 6 # 1.3 Assessment Plan Components Lab Test Flexible Assessments End semester/ Makeup

--- --- --- --- ---

(4- 5 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes Weightage 0.3 0.2 0.5

--- --- --- ---

Typology of Applying; Analyzing. Applying; Analyzing. Applying; Analyzing; Evaluating.

questions Evaluating.

Pattern Answer all the questions. Maximum Assignment: Data Extraction, Batch possessing and handling real time Answer all the questions. Maximum

marks 50.

marks 30. processing

Schedule As per academic Assignment submission: November As per academic calendar.

calendar. 2024

Topics covered HDFS, SQOOP, HIVE,

Map-Reduce program, Comprehensive examination

Spark, Data Frames and covering the full syllabus.

Data Streaming #### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

Lab1 Installing and configuring MySQL CO1

Lab2 Write script to handle data in MySQL, Shell commands for HDFS CO1

Lab3 Data ingestion using SQOOP CO2

Lab4 Data Analysis using HIVE CO2

Lab5 Introduction to Map-Reduce program CO2

Lab6 Experiment with Map-Reduce programs CO2

IT1 Internal lab test CO1, CO2

Lab7 Introduction to Spark Program CO3

Lab8 Data handling with Spark RDD CO3

Lab9 Data Analysis with Spark Data Frames CO3

Lab10 Data Streaming applications with Spark Structured Streams CO3

Lab11 Develop batch processing application CO4

Lab12 Develop real time processing application CO4 #### 1.5 References- 1. Big Data: Principles and best practices of scalable real-time data systems Nathan Marz and James Warren. Manning Publisher.

- 2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale Tom White, O'Reilly Publication 4th Edition.

- 3. Spark: The Definitive Guide: Big Data Processing Made Simple Bill Chambers, Matei Zaharia, O'Reilly Publication 1st Edition.# 1.6 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.# 1.7 Course Timetable 1 st Semester Big Data Analytics Room: LG1 LH 3 Lab: Data Science Lab

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE

WED ABD LAB

THU

FRI

SAT #### 1.8 Assessment Plan Cos Marks & weightage

--- --- --- --- --- ---

CO No. CO Name Lab Test Assignment End Semester CO wise

(Max. 30) (Max. 20) (Max. 50) Weightage

CO1 Use data extraction tools to ingest various types of data 5 5 5 0.15

into big data systems.

CO2 Experiment with different tools and frameworks of 10 5 15 0.3

Hadoop eco-system.

CO3 Experiment with Spark Engine to process real-time data. 10 5 20 0.35

Design applications to handle batch and streaming data

CO4 5 5 10 0..2

using Hadoop and Spark tools.

Marks (weightage) 0.3 0.2 0.5 1.0 - In-semester Assessment is considered as the Internal Assessment (IA) in each subject for 50 marks, which includes the performances in class / tutorial participation, assignment work, lab work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for each lab subject is conducted for a maximum of 50.

- End-semester mark for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in a subject.### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. Tools (TLP) Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- --- ---

No.

1 Sessional 0.3 2 • • Performance is measured using sessional attainment level. Reference: question paper and answer scheme.

• Each test is assessed for a maximum of 30 marks.

2 Assignments 0.2 - • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 20 marks.

3 ESE 0.5 1 • Performance is measured using ESE attainment level.

• Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 50 mark. # 1.10 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y CO3 Y

--- --- --- ---

CO4 Y

Average Articulation Level \* \* \* #### \*\*Note: Enter correlation levels 1, 2 or 3 as defined below for both CO-PO mapping:\*\*- 1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High)

- If there is no correlation, apply "-"# \*\*Master of Engineering - ME (Big Data Analytics)\*\* Course Name : Fundamentals of Machine Learning

--- --- ---

Course Code : BDA 5103

Academic Year : 2024 - 25

Semester : I

Name of the Course Coordinator : Dr. Arockiaraj S

Name of the Program Coordinator : Dr. Prathviraj N #### \*\*Course File\*\* Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date 1. Course Plan 5

--- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping 6

1.3 Assessment Plan 7

1.4 Lesson Plan 8

1.5 References 10

1.6 Other Resources (Online, Text, Multimedia, etc.) 10

1.7 Course Timetable 11

1.8 Assessment Plan 12

1.9 Assessment Details 13

1.10 Course Articulation Matrix 14 ## Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences. ## Program Outcomes (POs)By the end of the postgraduate program in \*\*ME (Big Data Analytics)\*\*, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document.

PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. ## 1. Course Plan#### 1.1 Primary Information Course Name : Fundamentals of Machine Learning [BDA 5103]

--- --- ---

L-T-P-C : 3-0-0-3

Contact Hours : 36 Hours

Pre-requisite : Basic Programming with Python

Core/ PE/OE : Core ### 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

CO1 Apply different types of supervised and unsupervised 8 PO3 3

machine learning algorithms to practical problems.

CO2 Analyse different types of machine learning paradigms. 12 PO4 4

CO3 Evaluate the performance of machine learning algorithms. 16 PO5 5 #### 1.3 Assessment Plan Components Mid semester Flexible Assessments End semester/ Makeup

--- --- --- ---

(2 – 3 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.3 0.2 0.5

Typology of Applying; Analyzing and Evaluating. Applying; Analyzing. Applying; Analyzing;

questions Evaluating. Evaluating.

Answer all 5 questions of 10 marks each. Assignment: Solving problems by Answer all 10 full questions of 10 marks each. Each question

Pattern Each question may have 2 to 3 parts of applying, analyzing and evaluating may have 2 to 3 parts of

3/4/5/6/7 marks. Generative AI use cases. 3/4/5/6/7 marks.

Schedule As per academic calendar. Assignment submission: November As per academic calendar.

2024

Introduction to Machine Learning; Decision Generative AI use cases, project Comprehensive examination

Topics Trees- Linear Model: K-nearest Neighbours lifecycle, and model pre-training - Fine covering the full syllabus.

covered Algorithm- Cross-validation - Dimension tuning and evaluating large language Students are expected to

Reduction. models. answer all questions. #### 1.4 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

L1 Overview of Supervised (regression and classification), unsupervised (clustering and dimensionality reduction), semi-supervised, and reinforcement learning with practical examples. CO1

L2 Machine learning nomenclature: raw data, types of features and outputs, feature vector. CO1

L3 Decision tree model of learning. Classification and regression using decision trees. CO1

L4 Splitting criteria: entropy, information gain, Gini impurity. CO1

L5 Splitting criteria: entropy, information gain, Gini impurity. CO1

L6 Splitting criteria: Gini impurity. CO1

L7 Overfitting in decision trees. CO1

L8 Pruning in decision trees. CO1

L9 Linear regression: model, estimation, and interpretation of coefficients. CO2

L10 Introduction to bias/variance trade-off. CO2

L11 Regularized linear regression. CO2

L12 K-nearest neighbours algorithm. CO2

L13 K-nearest neighbours algorithm. CO2 L14 Cross-validation CO2

--- --- ---

L15 Dimension reduction using principal component analysis (PCA) CO2

L16 Dimension reduction using principal component analysis (PCA) CO2

L17 Dimension reduction using multidimensional scaling CO2

L18 Dimension reduction using multidimensional scaling CO2

L19 Dimension reduction using t-SNE (t-distributed Stochastic Neighbour Embedding). CO2

L20 Dimension reduction using t-SNE (t-distributed Stochastic Neighbour Embedding). CO2

IT1 Internal test 1 CO1 & CO2

L21 Bagging: classification using random forest. CO3

L22 Boosting. CO3

L23 Probabilistic modelling of data using parameters - Introduction to maximum likelihood estimation CO3

(MLE) of parameters.

L24 Probabilistic modelling of data using parameters - Introduction to maximum likelihood estimation CO3

(MLE) of parameters.

L25 Naive Bayes model for classification. CO3

L26 Naive Bayes model for classification. CO3

L27 Logistic regression for binary classification. CO3

L28 Logistic regression for binary classification. CO3

L29 Model performance metrics. CO3

L30 Cross-validation. CO3

L31 Generative AI use cases. CO3 L32 Generative AI use cases. CO3

--- --- ---

L33 Project lifecycle, and model pre-training. CO3

L34 Project lifecycle, and model pre-training. CO3

L35 Fine-tuning and evaluating large language models. CO3

L36 Fine-tuning and evaluating large language models. CO3 ### 1.5 References- 1. Module: Introduction to Machine Learning (https://www.intel.com/content/www/us/en/developer/tools/oneapi/training/academicprogram/educators/intro-machine-learning-training-kit.html)

- 2. Module: Get started with AI on Azure (https://learn.microsoft.com/en-us/training/modules/get-started-ai-fundamentals/)

- 3. Module: Microsoft Azure AI Fundamentals: Get started with artificial intelligence (https://learn.microsoft.com/en-us/training/paths/getstarted-with-artificial-intelligence-on-azure/)

- 4. Learning path: Understand data science for machine learning (https://learn.microsoft.com/en-us/training/paths/understand-machinelearning/)

- 5. Module: Generative AI with Large Language Models (https://www.coursera.org/learn/generative-ai-with-llms)

- 6. Grokking Machine Learning, Luis G. Serrano, Manning Publications; 1st Edition, 2019 Online resource from Manning Publications available at https://www.manning.com/books/grokking-machine-learning

- 7. A Course in Machine Learning, Hal Daumé III Online resource available at http://ciml.info/

- 8. An Introduction to Statistical Learning: with Applications in Python (Springer Texts in Statistics), Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, and Jonathan Taylor, 1st Edition, 2023 – Online resource available at https://www.statlearning.com/#### 1.6 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.#### 1.7 Course Timetable 1 st Semester Big Data Analytics Lecture Hall:

--- --- --- --- --- --- --- --- ---

9-10 10-11 11-12 12-1 1-2 2-3 3-4 4-5

MON

TUE FML

WED

THU FML

FRI

SAT FML #### 1.8 Assessment Plan COs Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name Mid semester Assignment End Semester CO wise

(Max. 50) (Max. 20) (Max. 100) Weightage

CO1 Apply different types of supervised and unsupervised machine learning algorithms to practical problems. 25 5 20 0.30

CO2 Analyse different types of machine learning paradigms. 25 5 40 0.40

CO3 Evaluate the performance of machine learning algorithms. - 10 40 0.30

Marks (weightage) 0.3 0.2 0.5 1.0 Note:- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.

- End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.Weightage for CO1 = (mid semester marks for CO1 / 1.6666 + Assignment marks for CO1/1.0 + ESE marks for CO1 / 2)/100#### 1.9 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

• Performance is measured using internal test attainment level.

1 Internal Test 0.3 1 • Reference: question paper and answer scheme.

• Each internal test is assessed for a maximum of 50 marks and scaled down to 40

marks.

2 Assignments 0.2 2 • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 10 marks.

• Performance is measured using ESE attainment level.

3 End semester 0.5 1 • Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 100 marks and scaled down to 50 marks. #### 1.10 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y

CO3 Y

Average Articulation Level Y Y Y # Manipal School of Information Sciences (MSIS) \*\*Manipal Academy of Higher Education, Manipal Master of Engineering - ME (Big Data Analytics)\*\*#### \*\*Course File\*\* Course Name : Architecture of Big Data Systems

--- --- ---

Course Code : BDA 5102

Academic Year : 2024 – 2025

Semester : I

Name of the Course Coordinator : Mr. DEEPAK RAO B

Name of the Program Coordinator : Dr. PRATHVIRAJ N Signature of Program Coordinator Signature of Course Coordinator

--- ---

with Date with Date Course Plan 5 1.

--- --- ---

1.1 Primary Information 5

1.2 Course Outcomes (COs) Error! Bookmark not defined.

1.3 Course Content (Syllabus) Error! Bookmark not defined.

1.4 References Error! Bookmark not defined.

1.5 Other Resources (Online, Text, Multimedia, etc.) 11

1.6 Topic Learning Outcomes (TLOs) Error! Bookmark not defined.

1.7 Course Timetable Error! Bookmark not defined.

1.8 Assessment Plan Error! Bookmark not defined.

1.9 Assessment Details 13

1.10 Mapping of COs with POs Error! Bookmark not defined.

Assessment Details Error! Bookmark not defined. 2.

2.1 Student Details: Error! Bookmark not defined.

2.2 Assessment outcomes Error! Bookmark not defined.

2.3 Analysis of Assessment outcomes Error! Bookmark not defined.

2.4 Attainment of Course Outcomes (Direct) Error! Bookmark not defined. 2.5 Attainment of Course Outcomes (Indirect): Course End Survey (CES) Questionnaire Error! Bookmark not defined.

--- --- --- ---

2.6 Attainment of Course Outcomes (Indirect): Analysis Error! Bookmark not defined.

3. CO-PO Assessment Error! Bookmark not defined.

4. Observations and Comments Error! Bookmark not defined.

4.1 Observations from Course Coordinator based on the direct assessment Error! Bookmark not defined.

4.2 Comments/Suggestions by the Course Coordinator Error! Bookmark not defined. ## Program Education Objectives (PEOs)The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for \*\*ME (Big Data Analytics)\*\*, program are as follows. PEO No. Education Objective

--- ---

PEO 1 Develop in depth understanding of the key technologies in data engineering, data science and business analytics.

PEO 2 Practice problem analysis and decision-making using machine learning techniques.

PEO 3 Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences. ### Program Outcomes (POs)By the end of the postgraduate program in Big Data Analytics, graduates will be able to: PO1 Independently carry out research /investigation and development work to solve practical problems.

--- ---

PO2 Write and present a substantial technical report/document. PO3 Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level

--- ---

higher than the requirements in the appropriate bachelor program.

PO4 Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed

knowledge of the underlying data.

PO5 Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making. ### 1. Course Plan#### 1.1 Primary Information Course Name : Architecture of Big Data Systems

--- --- --- ---

L-T-P-C : 3-0-0-3

Contact Hours : 36 Hours

Pre-requisite : Programming with Python or Java #### 1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping CO At the end of this course, the student should be able to: No. of Contact Program Outcomes BL

--- --- --- --- ---

Hours (PO's)

CO1 Apply various techniques to examine different types of 12 PO3 3

data and understand lambda architecture.

CO2 Apply different tools and frameworks of Hadoop eco system 9 PO4 4

CO3 Apply Spark engine to process real-time data. 9 PO4 4

CO4 Design applications to handle batch and streaming data 6 PO5 6

using Hadoop and Spark tools. #### 1.3 Assessment Plan 2. Components Mid Semester Flexible Assessments End semester/ Makeup

--- --- --- ---

(2 – 3 in number) examination

Duration 90 minutes To be decided by the faculty. 180 minutes

Weightage 30% 20% 50%

Typology of questions Applying; Analyzing and Applying; Analyzing. Evaluating. Applying; Analyzing; Evaluating.

Evaluating

Pattern Answer all 5 questions of 10 Quiz / Test: 10% Assignment 1: (Big Data and weightage) Answer all 10 full questions of

Lambda Architecture, 5%

marks each. Each 10 marks each. Each question

question may have may have 2 to 3 parts of

2 to 3 parts of Assignment 2: (Spark 3/4/5/6/7 marks.

Streaming, 5% weightage)

3/4/5/6/7 marks.

Schedule As per academic Quiz / Test: September 2024 October 2024 As per academic calendar.

Assignment 1:

calendar. Assignment 2: November 2024 Introduction – Big Comprehensive examination

--- --- --- ---

Topics covered Data, Lambda covering the full syllabus.

Architecture, Batch Students are expected to

Processing, Spark answer all questions. Components Internal Test 1 Internal Test 2 Flexible Assessments End semester/ Makeup

--- --- --- --- ---

(2 – 3 in number) examination

Duration 90 minutes 90 minutes To be decided by the faculty. 180 minutes

Weightage 0.2 0.2 0.1 0.5

Typology of questions Applying; Applying; Applying; Analyzing. Applying; Analyzing;

Analyzing. Analyzing. Evaluating.

Pattern Answer all 5 Answer all 5

questions of 10 questions of 10 Assignment: (Big Data and Answer all 10 full questions of

marks each. Each marks each. Each Lambda Architecture, Hadoop ad 10 marks each. Each question

question may have 2 question may have Spark Assignments) may have 2 to 3 parts of

to 3 parts of 2 to 3 parts of 3/4/5/6/7 marks.

3/4/5/6/7 marks. 3/4/5/6/7 marks. Schedule As per academic As per academic Assignment submission: As per academic calendar.

--- --- --- --- ---

calendar. calendar. November 2023

Topics covered Introduction – Big Hadoop eco Comprehensive examination

Data, Lambda system, Realtime covering the full syllabus.

Architecture, Batch data processing Students are expected to answer

Processing using Spark. all questions. #### 2.1 Lesson Plan L. No. TOPICS Course Outcome Addressed

--- --- ---

L0 Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books ---

L1 Definition of Big Data and its sources CO1

L2 Characteristics of Big Data CO1

L3 Challenges of Big Data CO1

L4 Drawbacks of traditional data handling systems CO1

L5 Introduction to Lambda architecture CO1

L6 Requirements and responsibilities of batch layer CO1

L7 Requirements and responsibilities of service layer CO1 L8 Requirements and responsibilities of speed layer CO1

--- --- ---

L9 Introduction to Hadoop CO1

L10 Architecture of HDFS CO1

L11 Name Node and Data Nodes CO1

L12 How to ensure high availability of data and services CO1

L13 Introduction to YARN CO2

L14 Components of YARN CO2

L15 Introduction to Hadoop map-reduce CO2

L16 Responsibilities of Map task CO2

L17 Responsibilities of Reduce task CO2

L18 Types of failures in map reduce jobs CO2

L19 Recovery from failures CO2

L20 Writing map reduce programs CO2

L21 Writing map reduce programs CO2

L22 Introduction to Spark Engine CO3

L23 Difference between Spark and Hadoop CO3

L24 Architecture of Spark CO3

Mid Semester Evaluation CO1, CO2, CO3

L25 Different components and responsibilities of Spark CO3

L26 Different stages of running jobs in Spark CO3

L27 Actions and Transformations CO3 L28 RDDs and Data Frames CO3

--- --- ---

L29 Introduction to Data Streaming CO3

L30 Different models in Data Streaming CO3

L31 Challenges in data streaming CO4

L32 Developing programs for real time data handling CO4

L33 Developing programs for real time data handling CO4

L34 Develop programs for machine learning CO4

L35 Developing programs using Hadoop tools CO4

L36 Develop programs using Spark CO4 #### 2.2 References- 1. Big Data: Principles and best practices of scalable real-time data systems Nathan Marz and James Warren. Manning Publisher

- 2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale Tom White, O'Reilly Publication 4th Edition.

- 3. Spark: The Definitive Guide: Big Data Processing Made Simple Bill Chambers, Matei Zaharia, O'Reilly Publication 1st Edition.#### 2.3 Other Resources (Online, Text, Multimedia, etc.)- 1. Web Resources: Blog, Online tools and cloud resources.

- 2. Journal Articles.#### 2.4 Course Timetable 1st Semester BDA Room: LG1 LH8 Lab: Data Science

--- --- --- --- --- --- --- --- --- --- --- --- ---

9 – 10 10 - 11 11 – 12 12 2 - 3 3 - 4 4 - 5

MON

TUE ABD

WED

THU ABD

FRI

SAT ABD #### 2.5 Assessment Plan Cos Marks & Weightage

--- --- --- --- --- ---

CO No. CO Name IT-1 Assignment End Semester CO wise

(Max. 50) (Max. 10) (Max. 100) Weightage

CO1 Apply various techniques to examine different types 20 5 30 0.32

of data and understand lambda architecture.

CO2 Apply different tools and frameworks of Hadoop 20 10 30 0.35

eco-system CO3 Apply Spark engine to process real-time data. 10 5 20 0.21

--- --- --- --- --- ---

CO4 Design applications to handle batch and streaming - 20 0.12

data using Hadoop and Spark tools.

Marks (weightage) 0.3 0.2 0.5 1.0 - In-semester Assessment is considered as the Internal Assessment (IA) in each subject for 50 marks, which includes the performances in class / tutorial participation, assignment work, lab work, class tests, mid-term tests, quizzes etc.

- End-semester examination (ESE) for each theory subject is conducted for a maximum of 100 and the same will be scaled down to 50.

- End-semester mark for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in a subject.#### 2.6 Assessment DetailsThe assessment tools to be used for the Current Academic Year (CAY) are as follows: SI. No. Tools (TLP) Weightage Frequency Details of Measurement (Weightage/Rubrics/Duration, etc.)

--- --- --- --- ---

1 Mid Semester 0.3 • 1 Performance is measured using sessional attainment level.

• Reference: question paper and answer scheme. • Mid semester is assessed for a maximum of 50 marks and scaled down to 30

--- --- --- --- --- --- ---

marks

2 Assignments 0.2 1 • Performance is measured using assignments/quiz attainment level.

• Assignments/quiz are evaluated for a maximum of 20 marks.

• Performance is measured using ESE attainment level.

3 ESE 0.5 1 • Reference: question paper and answer scheme.

• ESE is assessed for a maximum of 100 marks and scaled down to 50 marks. #### 2.7 Course Articulation Matrix CO PO1 PO2 PO3 PO4 PO5

--- --- --- --- --- ---

CO1 Y

CO2 Y

CO3 Y

CO4 Y

Average Articulation Level \* \* \*