

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

Programme Name/s : Cloud Computing and Big Data
Programme Code : BD
Semester : Sixth
Course Title : SOFTWARE ENGINEERING AND TESTING FOR BIG DATA
Course Code : 316324

I. RATIONALE

This course provides a structured approach to software engineering, integrating Big Data principles for designing, estimating, testing, and quality management in large-scale software development. It enables students to handle massive datasets, optimize performance, ensure data integrity, and enhance software reliability using data-driven testing strategies. By blending domain knowledge, programming skills, and Big Data testing techniques, learners will be equipped to develop scalable, high-quality software solutions.

II. INDUSTRY / EMPLOYER EXPECTED OUTCOME

The aim of this course is to help the student to attain the following industry relevant outcome through various teaching learning experiences:

Apply software engineering principles and Big Data testing methodologies to develop scalable, high-performance, and reliable software products capable of handling large-scale data processing and analytics effectively.

III. COURSE LEVEL LEARNING OUTCOMES (COS)

Students will be able to achieve & demonstrate the following COs on completion of course based learning

- CO1 - Identify relevant software process model for software development.
- CO2 - Use appropriate software modelling principles to design software.
- CO3 - Apply project management techniques in software development.
- CO4 - Apply different software testing methodologies to ensure its quality.
- CO5 - Apply software testing processes and tools for big data applications.

IV. TEACHING-LEARNING & ASSESSMENT SCHEME

Course Code	Course Title	Abbr	Course Category/s	Learning Scheme						Credits	Assessment Scheme												Total Marks			
				Actual Contact Hrs./ Week	SLH			NLH			Paper Duration	Theory				Based on LL & TL				Based on SL						
																Practical										
												CL	TL	LL	FA-TH	SA-TH	Total		FA-PR		SA-PR			SLA		
Max	Max	Max	Min	Max	Min	Max	Min	Max	Min																	
316324	SOFTWARE ENGINEERING AND TESTING FOR BIG DATA	SEB	DSC	3	-	4	1	8	4	3	30	70	100	40	25	10	-	-	25	10	150					

Total IKS Hrs for Sem. : Hrs

Abbreviations: CL- Classroom Learning, TL- Tutorial Learning, LL-Laboratory Learning, SLH-Self Learning Hours, NLH-Notional Learning Hours, FA - Formative Assessment, SA -Summative assessment, IKS - Indian Knowledge System, SLA - Self Learning Assessment

Legends: @ Internal Assessment, # External Assessment, *# On Line Examination, @\$ Internal Online Examination

Note :

1. FA-TH represents average of two class tests of 30 marks each conducted during the semester.
2. If candidate is not securing minimum passing marks in FA-PR of any course then the candidate shall be declared as "Detained" in that semester.
3. If candidate is not securing minimum passing marks in SLA of any course then the candidate shall be declared as fail and will have to repeat and resubmit SLA work.
4. Notional Learning hours for the semester are (CL+LL+TL+SL)hrs.* 15 Weeks
5. 1 credit is equivalent to 30 Notional hrs.

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

6. * Self learning hours shall not be reflected in the Time Table.
 7. * Self learning includes micro project / assignment / other activities.

V. THEORY LEARNING OUTCOMES AND ALIGNED COURSE CONTENT

Sr.No	Theory Learning Outcomes (TLO's) aligned to CO's.	Learning content mapped with Theory Learning Outcomes (TLO's) and CO's.	Suggested Learning Pedagogies.
1	TLO 1.1 Explain different types and characteristics of software. TLO 1.2 Describe software engineering layered technology and process framework. TLO 1.3 State software engineering principles for requirement engineering. TLO 1.4 Explain the features of different software development process model for the given problem statement. TLO 1.5 Apply agile development process with justification.	Unit - I Basics of Software Engineering 1.1 Software, software engineering as layered approach, characteristics of software, types of software 1.2 Software development framework: Software generic process framework activities and umbrella activities 1.3 Software engineering core principles, communication practices, planning practices, modelling practices, construction practices, software deployment practices 1.4 Software Development Life Cycle 1.5 Prescriptive process models: Waterfall model, incremental model, RAD model, prototyping model, spiral model 1.6 Agile software development: Agile process, and its importance, extreme programming, scrum 1.7 Selection criteria for software process model	Presentations Video Demonstrations Lecture Using Chalk-Board
2	TLO 2.1 Perform requirement analysis for given software development problem. TLO 2.2 Prepare use case diagram for given scenario. TLO 2.3 Prepare SRS for the given problem. TLO 2.4 Convert analysis model into requirement model. TLO 2.5 Apply the specified design feature for requirements software modeling. TLO 2.6 Represent the specified problem in the given design notation.	Unit - II Software Requirement, Modeling and Design 2.1 Requirement engineering: Requirement engineering task, types of requirement, developing use-case 2.2 SRS (Software Requirements Specifications): Need of SRS, format and it's characteristics 2.3 Translating requirement model into design model 2.4 Design modelling: Fundamental design concepts - abstraction, information hiding, patterns, modularity, concurrency, verification, aesthetics 2.5 Design notations: Data flow diagram (DFD), structured flowcharts	Presentations Video Demonstrations Lecture Using Chalk-Board
3	TLO 3.1 Explain 4 P's of management spectrum. TLO 3.2 Estimate the size of the software product using the given method. TLO 3.3 Evaluate the cost of the given software using COCOMO model. TLO 3.4 Describe the RMMM strategy for the given problem. TLO 3.5 Use various scheduling techniques for the given project. TLO 3.6 Prepare the Timeline chart / Gantt chart to track progress of the given project.	Unit - III Software Project Management 3.1 The management spectrum- 4P's 3.2 Metrics for size estimation: Line of code (LoC), function points(FP) 3.3 Project cost estimation using COCOMO (Constructive Cost Model), COCOMO II 3.4 Define risk, types of risk, RMMM strategy 3.5 Project scheduling: Basic principle, scheduling techniques - CPM, PERT 3.6 Project tracking: Timeline charts, Gantt charts	Video Demonstrations Presentations Lecture Using Chalk-Board
4	TLO 4.1 Describe the objectives and importance of software	Unit - IV Software Testing and Quality Management	Presentations Video

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

Sr.No	Theory Learning Outcomes (TLO's) aligned to CO's.	Learning content mapped with Theory Learning Outcomes (TLO's) and CO's.	Suggested Learning Pedagogies.
	testing. TLO 4.2 Explain verification and validation processes in software testing. TLO 4.3 Explain different levels of testing: Unit testing, integration testing, system testing, and acceptance testing TLO 4.4 Elaborate the importance of test planning and how to prepare a test plan. TLO 4.5 Explain Defect management process.	4.1 Fundamentals of Software Testing: Software testing objectives, Software Testing Life Cycle (STLC), Failure, fault, error, defect, bug, Definition, Test Cases - when to start and stop testing 4.2 Quality Assurance and Control: Quality assurance, quality control, verification, and validation. 4.3 Testing Methodologies: Static and dynamic testing, White box testing vs. black box testing, Unit testing, integration testing, system testing, acceptance testing, Manual testing vs. automation testing 4.4 Test and Defect Management: Test planning: Preparing a test plan, Test infrastructure management, Executing test cases, preparing test summary reports 4.5 Defect management: Definition and types of defects, defect life cycle, defect template	Demonstrations Lecture Using Chalk-Board
5	TLO 5.1 Explain the importance of software testing for big data. TLO 5.2 Write methods to apply different software testing techniques as applicable to big data. TLO 5.3 Explain big data testing process. TLO 5.4 Explain the working of given software testing tools for big data.	Unit - V Software Testing for Big Data 5.1 Introduction to Big data testing, Importance of Big data Testing, Challenges of Testing with Big Data 5.2 Testing Techniques for Big Data :Functional & Non Functional 5.3 Big Data testing process: Data injection Testing, Data Processing Testing, Data Storage Testing, Data Validation Testing, Performance Testing, Security & Compliance testing. 5.4 Big Data Testing Tools	Presentations Video Demonstrations Lecture Using Chalk-Board

VI. LABORATORY LEARNING OUTCOME AND ALIGNED PRACTICAL / TUTORIAL EXPERIENCES.

Practical / Tutorial / Laboratory Learning Outcome (LLO)	Sr No	Laboratory Experiment / Practical Titles / Tutorial Titles	Number of hrs.	Relevant COs
LLO 1.1 Identify problem statements and define project scope for cloud-based big data applications.	1	Define a Problem Statement and Scope for a Cloud-based Big Data Project	2	CO1
LLO 2.1 Select a software process model suited for cloud-based development.	2	Select Appropriate Software Process Model for Cloud-based Big Data Project	2	CO1
LLO 3.1 Document requirements for cloud-hosted big data projects.	3	Document Functional and Non-Functional Requirements for Cloud-based Big Data Project	2	CO2
LLO 4.1 Create an SRS document based on standard industry formats.	4	Prepare Software Requirement Specification document for Cloud-hosted Big Data Application	2	CO2
LLO 5.1 Design use case diagram to model cloud-based data processing.	5	Create Use Case Diagram for Cloud Data Processing Workflow	2	CO2
LLO 6.1 Represent an ETL (Extract, Transform, Load) process in a Cloud environment using an activity diagram.	6	Draw Activity Diagram for Big Data ETL (Extract, Transform, Load) workflow in Cloud environment	2	CO2
LLO 7.1 Construct a decision table to represent the routing logic for diverse big data types.	7	Design Decision Table for Cloud-based Data Routing	2	CO2

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

Practical / Tutorial / Laboratory Learning Outcome (LLO)	Sr No	Laboratory Experiment / Practical Titles / Tutorial Titles	Number of hrs.	Relevant COs
LLO 8.1 Develop Data Flow Diagrams (DFDs) to visualize the movement of big data through various cloud components.	8	Create Data Flow Diagrams (DFDs) for Big Data Flow in Cloud Architecture	2	CO2
LLO 9.1 Construct a class diagram to define and organize big data entities.	9	Create a class diagram for big data entities in Cloud environment	2	CO2
LLO 10.1 Design a sequence diagram to illustrate the interaction flow of big data requests between cloud components.	10	Draw Sequence diagram for Big Data request handling in Cloud environment	2	CO2
LLO 11.1 Design a collaboration diagram to represent interactions between cloud Microservices in big data processing.	11	Draw a collaboration diagram for Cloud-based Big Data microservices interaction	2	CO2
LLO 12.1 Estimate project size using function point analysis.	12	Perform function point analysis to estimate the size of a Cloud-based Big Data application by identifying inputs, outputs, files, and interfaces	2	CO3
LLO 13.1 Use the COCOMO model to estimate the development cost, effort and time.	13	Apply the COCOMO model to estimate the development cost, effort, and time for a Cloud-based Big Data project	2	CO3
LLO 14.1 Prepare an RMMM (Risk Mitigation, Monitoring, and Management) plan to address and manage potential risks.	14	Prepare an RMMM (Risk Mitigation, Monitoring, and Management) plan to mitigate, monitor, and manage potential risks in a Cloud-based Big Data project	2	CO3
LLO 15.1 Use the Critical Path Method (CPM) technique to schedule tasks to determine the minimum project completion time.	15	Apply the Critical Path Method (CPM) to schedule tasks and determine the minimum completion time for deploying a Cloud-based Big Data project	2	CO3
LLO 16.1 Use the Program Evaluation and Review Technique (PERT) technique to estimate task durations to identify critical project activities.	16	Estimate Task Durations for Big Data Processing in a Cloud Environment using the Program Evaluation and Review Technique (PERT) to identify critical project activities	2	CO3
LLO 17.1 Develop a Gantt chart to plan, schedule and track the progress of tasks.	17	Create a Gantt Chart to plan, schedule, and track the progress of tasks in a Cloud-based Big Data project	2	CO3
LLO 18.1 Design functional test cases to validate system functionality as per requirements.	18	Execute functional test cases for Cloud-based Big Data operations	2	CO4 CO5
LLO 19.1 Develop GUI test cases to verify the usability, responsiveness, and correctness of user interfaces.	19	Execute GUI test cases to evaluate the usability, responsiveness, and accuracy of User Interfaces in Cloud-based Big Data Applications	2	CO4 CO5
LLO 20.1 Prepare a test plan outlining strategy, resources, and schedule for testing cloud microservices handling big data.	20	Prepare a test plan for testing Cloud Microservices managing Big Data operations	2	CO4 CO5
LLO 21.1 Implement automated GUI test cases.	21	Execute automated GUI test cases for Cloud-based Big Data Application using Selenium	2	CO4 CO5
LLO 22.1 Execute unit test cases for individual components of a cloud-based big data project.	22	Execute unit test cases for individual functions/modules in a cloud-based big data application using Python's built-in unittest framework	2	CO4 CO5
LLO 23.1 Perform Integration Testing to validate the seamless interaction and data flow between interconnected software components.	23	Perform integration testing to validate the seamless interaction and data flow between different cloud-based microservices and big data APIs	2	CO4 CO5

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

Practical / Tutorial / Laboratory Learning Outcome (LLO)	Sr No	Laboratory Experiment / Practical Titles / Tutorial Titles	Number of hrs.	Relevant COs
LLO 24.1 Apply load testing to simulate multiple user requests to analyze the system's performance under heavy load.	24	Apply Load Testing on Cloud-based Big Data Systems using JMeter	2	CO4 CO5
LLO 25.1 Execute Security Testing to identify vulnerabilities.	25	Perform basic security testing using open-source tools for a Cloud-based Big Data Application.	2	CO4 CO5
LLO 26.1 Test Big Data Storage and Retrieval Performance in a Cloud Environment.	26	Evaluate database performance by measuring query execution times in cloud environments	2	CO4 CO5
LLO 27.1 Perform Data Validation Testing to ensure accuracy, consistency, and completeness of data across different stages of processing in a software system.	27	Ensure data accuracy, consistency, and completeness across ingestion, processing, storage, and retrieval stages in a cloud-based big data pipeline	2	CO4 CO5
LLO 28.1 Perform regression tests to ensure software updates do not break existing functionality.	28	Execute Regression Test Cases for Cloud-based Big Data Systems	2	CO4 CO5
LLO 29.1 Use cloud-based logging and monitoring tools to debug system issues.	29	Analyze logs and debug errors in Cloud-based Big Data applications using tools like AWS CloudWatch, ELK Stack etc.	2	CO4 CO5
LLO 30.1 Deploy a basic cloud-hosted application.	30	Deploy a Simple Cloud-hosted Big Data Application	2	CO4 CO5

Note : Out of above suggestive LLOs -

- '*' Marked Practicals (LLOs) Are mandatory.
- Minimum 80% of above list of lab experiment are to be performed.
- Judicial mix of LLOs are to be performed to achieve desired outcomes.

VII. SUGGESTED MICRO PROJECT / ASSIGNMENT/ ACTIVITIES FOR SPECIFIC LEARNING / SKILLS DEVELOPMENT (SELF LEARNING)**Assignment**

- Explain the objectives and importance of software testing, the verification and validation processes, and the different levels of testing.
- Explain the Importance of Risk Management in Cloud Big Data Projects.
- Estimate size of software using any tool and risk involved in any food delivery system.
- Prepare test plan and defect report for calculator.
- Explain the different software development models their characteristics, and selection criteria.
- Explain and Model Big Data Workflows Using UML Diagrams.

Other

- Discuss paper titled "Case Study Based Software Engineering Project Development: State of Art" reference link: <https://arxiv.org/pdf/1306.2502>.

Micro project

- Build a basic cloud-hosted system to process a sample student dataset and display performance statistics.
- Process sample weather data and display average temperature and rainfall charts.
- Create a mock feedback form for a cloud application that collects and displays feedback data.
- Design a cloud-based tool to upload sample sales data and generate total sales reports.
- Simulate complaint submission, storage, and status tracking using a cloud-hosted form.

Note :

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

- Above is just a suggestive list of microprojects and assignments; faculty must prepare their own bank of microprojects, assignments, and activities in a similar way.
- The faculty must allocate judicious mix of tasks, considering the weaknesses and / strengths of the student in acquiring the desired skills.
- If a microproject is assigned, it is expected to be completed as a group activity.
- SLA marks shall be awarded as per the continuous assessment record.
- For courses with no SLA component the list of suggestive microprojects / assignments/ activities are optional, faculty may encourage students to perform these tasks for enhanced learning experiences.
- If the course does not have associated SLA component, above suggestive listings is applicable to Tutorials and maybe considered for FA-PR evaluations.

VIII. LABORATORY EQUIPMENT / INSTRUMENTS / TOOLS / SOFTWARE REQUIRED

Sr.No	Equipment Name with Broad Specifications	Relevant LLO Number
1	Google Doc, LibreOffice, OpenOffice	1,2,3,4,7,20
2	Open-source project management tools: Jira, Redmine, Taiga, Trac, MantisBT	14,15,16,17,19
3	draw.io tool: Flow chart maker, UML diagram, use-case diagrams.	5,6,8,9,10,11
4	Windows 10 or Higher, Processor i3 or higher, HDD/SDD 500GB, RAM 8GB (recommended 16GB)	All

IX. SUGGESTED WEIGHTAGE TO LEARNING EFFORTS & ASSESSMENT PURPOSE (Specification Table)

Sr.No	Unit	Unit Title	Aligned COs	Learning Hours	R-Level	U-Level	A-Level	Total Marks
1	I	Basics of Software Engineering	CO1	6	4	4	4	12
2	II	Software Requirement, Modeling and Design	CO2	10	2	6	6	14
3	III	Software Project Management	CO3	10	2	2	12	16
4	IV	Software Testing and Quality Management	CO4	12	2	4	12	18
5	V	Software Testing for Big Data	CO5	7	2	4	4	10
Grand Total				45	12	20	38	70

X. ASSESSMENT METHODOLOGIES/TOOLS**Formative assessment (Assessment for Learning)**

- For formative assessment of laboratory learning 25 marks.
- Each practical will be assessed considering 60% weightage to process, 40% weightage to product.
- For theory two offline unit tests of 30 marks and average of two unit test marks will be considered for out of 30 marks.

Summative Assessment (Assessment of Learning)

- End semester assessment is of 70 marks.
- End Semester Examination, SLA evaluation, Viva-Voce.

XI. SUGGESTED COS - POS MATRIX FORM

Course Outcomes (COs)	Programme Outcomes (POs)	Programme Specific Outcomes* (PSOs)
-----------------------	--------------------------	-------------------------------------

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

	PO-1 Basic and Discipline Specific Knowledge	PO-2 Problem Analysis	PO-3 Design/ Development of Solutions	PO-4 Engineering Tools	PO-5 Engineering Practices for Society, Sustainability and Environment	PO-6 Project Management	PO-7 Life Long Learning	PSO-1	PSO-2	PSO-3
CO1	1	2	1	2	1	-	2			
CO2	2	2	1	2	-	-	2			
CO3	1	2	2	3	-	2	2			
CO4	2	2	2	2	1	2	2			
CO5	2	2	2	3	1	1	2			

Legends :- High:03, Medium:02,Low:01, No Mapping: -

*PSOs are to be formulated at institute level

XII. SUGGESTED LEARNING MATERIALS / BOOKS

Sr.No	Author	Title	Publisher with ISBN Number
1	Roger S. Pressman & Bruce R. Maxim	Software Engineering: A practitioner's approach	McGraw Hill Higher Education, New Delhi, (Ninth Edition) ISBN 93-5532-504-5
2	Srinivasan Desikan, Gopalaswamy Ramesh	Software Testing: Principles and Practices	PEARSON Publisher: Pearson India 2007, ISBN: 978-81-7758-121-8,
3	Richard Fairly	Software Engineering Concepts	McGraw Hill Education New Delhi -2001, ISBN-13: 9780074631218
4	Deepak Jain	Software Engineering: Principles and practices	Oxford University Press, New Delhi ISBN 9780195694840
5	Ron Patton	Software Testing	Sams Publishing; 2nd edition, 2005 ISBN: 0672327988
6	M. G. Limaye	Software Testing: Principles, Techniques and Tools	Tata McGraw Hill Education, New Delhi., 2009 ISBN 13: 9780070139909
7	Naresh Chauhan	Software Testing: Principles and Practices	Oxford University Press Noida. ISBN: 9780198061847
8	Yogesh Singh	Software Testing	Cambridge University Press, Cambridge, 2021 ISBN: 9781107012967

XIII . LEARNING WEBSITES & PORTALS

Sr.No	Link / Portal	Description
1	https://www.coursera.org/specializations/software-engineering	Software Engineering skill and techniques being used in Software Development Life Cycle (SDLC)
2	www.tutorialspoint.com/software_engineering/	Software Engineering Tutorial
3	https://www.tutorialspoint.com/software_testing/index.htm	Software Testing Tutorial
4	https://www.coursera.org/learn/foundations-of-software-testing-and-validation	Describe various types of software testing and identify their advantages and disadvantages.
5	https://nptel.ac.in/courses/106101061	Software Engineering
6	https://nptel.ac.in/courses/106105150	Software Testing
7	https://infyspringboard.onwingspan.com/web/en/app/toc/lex_27806881903165866000_shared/overview	BigData Testing Methodology
8	https://testlio.com/blog/software-quality-big-data/	Big Data & Software Testing: A Complete Guide

Note :

SOFTWARE ENGINEERING AND TESTING FOR BIG DATA**Course Code : 316324**

Sr.No	Link / Portal	Description
Teachers are requested to check the creative common license status/financial implications of the suggested online educational resources before use by the students		

MSBTE Approval Dt. 04/09/2025**Semester - 6, K Scheme**