University of Mumbai

वेबसाईट – mu.ac.in
ईमेल - आयडी - dr.aams@fort.mu.ac.in
aams3@mu.ac.in

विद्याविषयक प्राधिकरणे
सभा आणि सेवा विभाग(ए.ए.एम.एस)
रूम नं. १२८ एम.जी.रोड, फोर्ट,
मुंबई – ४०० ०३२
टेलिफोन नं. – ०२२ – ६८३२००३३

(नॅक पुनर्मूल्यांकनाद्वारे ३.६५ (सी.जी.पी.ए.) सह अ++ श्रेणी
विद्यापीठ अनुदान आयोगाद्वारे श्रेणी १ विद्यापीठ दर्जा)

क्र.वि.प्रा.स.से./आयसीडी/२०२५-२६/३७ दिनांक : २७ मे, २०२५

परिपत्रक:-

सर्व प्राचार्य/संचालक, संलग्नित महाविद्यालये/संस्था, विद्यापीठ शैक्षणिक विभागांचे संचालक/ विभाग प्रमुख यांना कळविण्यात येते की, राष्ट्रीय शैक्षणिक धोरण २०२० च्या अमंलबजावणीच्या अनुषंगाने **शैक्षणिक वर्ष २०२५-२६** पासून **पदवी व पदव्युत्तर अभ्यासक्रम** विद्यापरिषदेच्या दिनांक २८ मार्च २०२५ व २० मे, २०२५ च्या बैठकीमध्ये मंजूर झालेले सर्व अभ्यासक्रम मुंबई विद्यापीठाच्या www.mu.ac.in या संकेत स्थळावर NEP २०२० या टॅब वर उपलब्ध करण्यात आलेले आहेत.

मुंबई – ४०० ०३२
२७ मे, २०२५

(डॉ. प्रसाद कारंडे)
कुलसचिव

Copy forwarded for information and necessary action to :-	
1	The Deputy Registrar, (Admissions, Enrolment, Eligibility and Migration Dept)(AEM), dr@eligi.mu.ac.in
2	The Deputy Registrar, Result unit, Vidyanagari drresults@exam.mu.ac.in
3	The Deputy Registrar, Marks and Certificate Unit,. Vidyanagari dr.verification@mu.ac.in
4	The Deputy Registrar, Appointment Unit, Vidyanagari dr.appointment@exam.mu.ac.in
5	The Deputy Registrar, CAP Unit, Vidyanagari cap.exam@mu.ac.in
6	The Deputy Registrar, College Affiliations & Development Department (CAD), deputyregistrar.uni@gmail.com
7	The Deputy Registrar, PRO, Fort, (Publication Section), Pro@mu.ac.in
8	The Deputy Registrar, Executive Authorities Section (EA) eau120@fort.mu.ac.in He is requested to treat this as action taken report on the concerned resolution adopted by the Academic Council referred to the above circular.
9	The Deputy Registrar, Research Administration & Promotion Cell (RAPC), rapc@mu.ac.in
10	The Deputy Registrar, Academic Appointments & Quality Assurance (AAQA) dy.registrar.tau.fort.mu.ac.in ar.tau@fort.mu.ac.in
11	The Deputy Registrar, College Teachers Approval Unit (CTA), concolsection@gmail.com
12	The Deputy Registrars, Finance & Accounts Section, fort draccounts@fort.mu.ac.in
13	The Deputy Registrar, Election Section, Fort drelection@election.mu.ac.in
14	The Assistant Registrar, Administrative Sub-Campus Thane, thanesubcampus@mu.ac.in
15	The Assistant Registrar, School of Engg. & Applied Sciences, Kalyan, ar.seask@mu.ac.in
16	The Assistant Registrar, Ratnagiri Sub-centre, Ratnagiri, ratnagirisubcentar@gmail.com
17	The Director, Centre for Distance and Online Education (CDOE), Vidyanagari, director@idol.mu.ac.in
18	Director, Innovation, Incubation and Linkages, Dr. Sachin Laddha pinkumanno@gmail.com
19	Director, Department of Lifelong Learning and Extension (DLLE), dlleuniversityofmumbai@gmail.com

Copy for information :-	
1	P.A to Hon'ble Vice-Chancellor, vice-chancellor@mu.ac.in
2	P.A to Pro-Vice-Chancellor pvc@fort.mu.ac.in
3	P.A to Registrar, registrar@fort.mu.ac.in
4	P.A to all Deans of all Faculties
5	P.A to Finance & Account Officers, (F & A.O), camu@accounts.mu.ac.in

To,

1	The Chairman, Board of Deans pvc@fort.mu.ac.in
2	**Faculty of Humanities,** **Offg. Dean** 1. Prof.Anil Singh Dranilsingh129@gmail.com **Offg. Associate Dean** 2. Prof.Manisha Karne mkarne@economics.mu.ac.in 3. Dr.Suchitra Naik Naiksuchitra27@gmail.com
	Faculty of Commerce & Management, **Offg. Dean,** 1 Prin.Ravindra Bambardekar principal@model-college.edu.in **Offg. Associate Dean** 2. Dr.Kavita Laghate kavitalaghate@jbims.mu.ac.in 3. Dr.Ravikant Balkrishna Sangurde Ravikant.s.@somaiya.edu 4. Prin.Kishori Bhagat kishoribhagat@rediffmail.com

	Faculty of Science & Technology **Offg. Dean** 1. Prof. Shivram Garje ssgarje@chem.mu.ac.in **Offg. Associate Dean** 2. Dr. Madhav R. Rajwade Madhavr64@gmail.com 3. Prin. Deven Shah sir.deven@gmail.com
	Faculty of Inter-Disciplinary Studies, **Offg. Dean** 1.Dr. Anil K. Singh aksingh@trcl.org.in **Offg. Associate Dean** 2.Prin.Chadrashekhar Ashok Chakradeo cachakradeo@gmail.com 3. Dr. Kunal Ingle drkunalingle@gmail.com
3	Chairman, Board of Studies,
4	The Director, Board of Examinations and Evaluation, dboee@exam.mu.ac.in
5	The Director, Board of Students Development, dsd@mu.ac.in DSW direcotr@dsw.mu.ac.in
6	The Director, Department of Information & Communication Technology, director.dict@mu.ac.in

AC – 20/05/2025
Item No. – 8.51 (N)

As Per NEP 2020

University of Mumbai

Program: B.Sc. (Data Science and Sports Studies)

Semester –III , IV, V, VI

Ref: GR dated 20th April, 2023 for Credit Structure of UG

(With effect from the academic year 2024-25 Progressively)

University of Mumbai

(As per NEP 2020)

Sr. No.	Heading	Particulars
1	**Title of program** **O:_____**	**B.SC. (Data Science and Sports Studies)**
2	**Exit Degree**	**U.G. Diploma in ____**
3	**Scheme of Examination** **R: _____**	NEP 40% Internal 60% External, Semester End Examination Individual Passing in Internal and External Examination
4	**Standards of Passing** **R: _____**	**40%**
5	**Credit Structure** **R. IMU-560C** **R. IMU-560D** **R. IMU-560E** **R. IMU-560F**	**Attached herewith**
6	**Semesters**	**Sem. III & IV**
7	**Program Academic Level**	**5.00**
8	**Pattern**	**Semester**
9	**Status**	**New**
10	**To be implemented from Academic Year**	**2025-26**

Sd/-	Sd/-	Sd/-	Sd/-
Sign of the BOS Chairman **Dr. Manoj N. Reddy** **Ad-hoc Board of Studies in Sports Science and Management**	**Sign of the Offg. Associate Dean** **Dr. C.A.Chakradeo** **Faculty of Interdisciplinary Studies**	**Sign of the Offg. Associate Dean** **Dr. Kunal Ingle** **Faculty of Interdisciplinary Studies**	**Sign of the Offg. Dean** **Prof. A. K. Singh** **Faculty of Interdisciplinary Studies**

Under Graduate Diploma in Data Science and Sports Studies

Credit Structure (Sem. III & IV)

	R. IMU-560C									
Level	**Semester**	**Major**		**Minor**	**OE**	**VSC, SEC (VSEC)**	**AEC, VEC, IKS**	**OJT, FP, CEP, CC,RP**	**Cum. Cr. / Sem.**	**Degree/ Cum. Cr.**
		Mandatory	**Electives**							
5.0	III	Data Warehousing and Multidimensional Modelling (4 credits), Software Engineering (4 credits)		4 To be picked from University Basket	2 To be picked from University Basket	Discrete Mathematics(2 credits),	AEC:2 To be picked from University Basket	FP: 2 CC:2 To be picked from University Basket	22	**UG Diploma 88**
	R. IMU-560D									
	IV	Descriptive Analytics (4 credits), Machine Learning (4 credits)		4 To be picked from University Basket	2 To be picked from University Basket	R Programming (2 credits)	AEC:2 To be picked from University Basket	CEP: 2 CC:2 To be picked from University Basket	22	
	Cum Cr.	28		10	12	6+6	8+4+2	8+4	88	
Exit option; Award of UG Diploma in Major and Minor with 80-88 credits and an additional 4 credits core NSQF course/ Internship OR Continuewith Major and Minor										

[Abbreviation - OE – Open Electives, VSC – Vocation Skill Course, SEC – Skill Enhancement Course, (VSEC), AEC – Ability Enhancement Course, VEC – Value Education Course, IKS – Indian Knowledge System, OJT – on Job Training, FP – Field Project, CEP – Community Engagement Project, CC – Co-Curricular, RP – Research Project]

Under Graduate Diploma in Data Science and Sports Studies

Credit Structure (Sem. V & VI)

Level	Semester	Major		Minor	OE	VSC, SEC (VSEC)	AEC, VEC, IKS	OJT, FP, CEP, CC,RP	Cum. Cr. / Sem.	Degree/ Cum. Cr.
		Mandatory	Electives							
	R. IMU-560E									
5.5	V	Predictive Analytics (4 credits), Big Data Analytics (4 credits) Project 1 (2 credits)	4	4 To be picked from University Basket	To be picked from University Basket	Distributed Systems (2 credits)		FP/CEP:2	22	**UG Degree 132**
	R. IMU-560F									
	VI	Social And Web Analytics (4 credits), Analytics for Industries (4 credits) Project 2 (2 credits),	4	4 To be picked from University Basket				OJT :4	22	
	Cum Cr.	48	8	18	12	8+6	8+4+2	8+6+4	132	
Exit option: Award of UG Degree in Major with 132 credits OR Continue with Major and Minor										

[Abbreviation - OE – Open Electives, VSC – Vocation Skill Course, SEC – Skill Enhancement Course, (VSEC), AEC – Ability Enhancement Course, VEC – Value Education Course, IKS – Indian Knowledge System, OJT – on Job Training, FP – Field Project, CEP – Community Engagement Project, CC – Co-Curricular, RP – Research Project]

Sem. - III

Semester III	
Subject Name	**Credit**
Major (Mandatory)	
Machine Learning	4
Foundation of AI	4
Minor	
To be picked from University Basket	2
To be picked from University Basket	2
OE - Open Electives	
To be picked from University Basket	2
VSC - Vocational Skills Course	
Internet and Web Programming	2
AEC - Ability Enhancement Course	
To be picked from University Basket	2
FP - Field Project	
To be picked from University Basket	2
CC - Co-curricular Course	
To be picked from University Basket	2

Syllabus
B.Sc. (Data Science and Sports Studies)
(Sem- III)

Semester III

Sr. No.	Subjects	Subject Type	Credits	Internal	External	Total
1	Data Warehousing and Multidimensional Modelling	Major	4	40	60	100
2	Software Engineering	Major	4	40	60	100
3	Internet and Web Programming	VSC	2	20	30	50
4	Pick from University Basket	Minor	2	20	30	50
5	Pick from University Basket	Minor	2	20	30	50
6	Pick from University Basket	OE	2	20	30	50
7	Pick from University Basket	AEC	2	20	30	50
8	Pick from University Basket	FP	2	20	30	50
9	Pick from University Basket	CC	2	20	30	50

BSc (Data Science and Sports Studies)
(Sem.- III)
Title of Paper: Data Warehousing and Multidimensional Modelling

Sr. No.	Heading	Particulars
1	**Description the course :**	• Understand the purpose and benefits of data warehousing in an organization. • Learn the process of designing and building a data warehouse. • Gain knowledge of multidimensional modeling techniques for representing complex business data.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: •Data Warehousing and Multidimensional Modeling is a comprehensive course that explores the fundamental concepts and methodologies involved in building and managing data warehouses. •The course emphasizes the design and implementation of multidimensional data models to facilitate efficient data analysis and decision-making.	
8	**Course Outcomes**: • Explore the principles of data integration and data quality in a data warehousing environment. • Understand the role of Extract, Transform, Load (ETL) processes in populating a data warehouse. • Develop skills in querying and analyzing data from a multidimensional data model. • Learn best practices for data warehouse administration and performance optimization.	

9	**Modules**:-	
	Module 1: 7 Hours An introduction to data warehousing, Data warehouse architectures, Two- layer architecture, Three-layer architecture, Data staging and ETL, Decision support applications, Data warehouse functional definitions.	
	Module 2: **7 Hours** Data warehouse & modeling, Data warehouse modeling, Types of data warehouse models, Data warehouse implementation, Normal system for metadata sharing, Data warehouse environment, Data warehouse and data mart, Operational data store, Data warehouse data modeling styles, OLTP data model considerations, Data warehouse usage, Data mart, Warehouse modeling approaches, Global data warehouse architecture, Independent data mart	

architecture, Approaches to implement the architecture, Data warehouse delivery process, Data warehouse modelling: Techniques and guidelines, ER modeling, The ER model, Normalization: First normal form, Normalization: Second normal form, Normalization: Third normal form, OLAP (Online Analytical Processing), Multi-dimensional data analysis, Drill-down and roll-up.

Module 3: 6 Hours

Requirements analysis: Base activities, Representation of the query as a cube, Requirements modeling, Measures, Candidate measures, Dimensions, The grain of a dimension, Granularity of a measure, Facts, Terminologies in a multi-dimension model, Advantage of multi-dimensional modelling, Multi-dimensional queries, Dimension entities and attributes, Dimension hierarchies, Aggregation levels, MD model structures, Star schema, Snowflake schema, Example for snowflake schema, Considerations star and snowflake, Hints, tips, and guidelines, Business-related facts, What is fact constellation schema? Applications for data warehouse, Fact identifiers, dimension keys and uniqueness, Identifiers in an MD model, Dimension roles, Requirements analysis, Solution validation, Base activities, Detailed dimension modeling, Why needed? Detailed dimension modeling– approach, Detailed dimension modeling– considerations.

Module 4: 6 Hours

R-OLAP, OLAP query against RDBMS (ROLAP), R-OLAP design techniques, OLAP, Fundamental OLAP analytical practices, Types of OLAP systems, Architecture of ROLAP, Star schema and star join queries, Star join support, Efficient star join, Dynamic bitmap index ANDing –DB2 UDB for distributed platforms, Star join: DB2 UDB for distributed platforms, Star join: Snowflake schema context, Collapse dimensions, Effects of dimension collapse, Dimension split, Redundant attributes in facts, Fact- oriented design techniques, Fact aggregation, Consolidated facts, Utilize cubing services to improve R-OLAP and M-OLAP performance, Cubing services architecture, Cubing services performance and scalability, Scalability, Cubing services security, Benefits of cubing services: Improved data mart performance, Data marts, Data marts: Materialized query tables, Materialized query table example, Performance with MQT refresh options, Implementation example: MQT, Multi-dimension cluster tables, MDC table example, Terminology-dimension, Terminology -slice, Terminology-cell, Block based index: Dynamic bitmap index ANDing, Query processing examples, MDC performance example, Example: Object size comparisons, Power of block prefetch coupled, Performance results for a query, Performance results from a query, Example: Index oring of block and RID indexes, Example: Joins with block index

Module 5 : 7 Hours

Introduction to MOLAP, MOLAP framework, M-OLAP architecture, OLAP schemes types, Cognos functional architecture, Cognos Transformer: Functions, Storage method, Hyperion, Hyperion data load, Hyperion: data load rules, Hyperion calculation storage types, MOLAP characteristics, ROLAP characteristics, Online data analysis.

Module 6: 8 Hours

Keeping history about dimensions, Customer attrition: Business requirements, Customer attrition: Exercise, Customer attrition: Base analysis, Customer attrition: Solution, Primary key change, The historical customer, Customer address: Business requirement, Customer address: Exercise, Customer address: Solution, Issues, Time normalization, Customer address: Time normalized, Neighborhoods: Business requirements, Neighborhoods: Exercise, Neighborhoods: Base solution, Neighborhoods: Review, Neighborhoods: HTK solution, Customer classes: Business requirement, Customer classes: Exercise, Customer classes: Proposed solution.

	Module 7 : 4 Hours Process architecture: Data warehouse, Contrasting OLTP and data warehousing environments, Common data warehouse tasks, Architectures of the data warehouse, Centralized process architecture, Distributed process architecture, Peer-to-Peer architecture, Data warehouse tools, 25 BEST data mining tools in 2020, 18 BEST reporting tools in 2020, IBM data warehouse case study: American Airlines.	
10	**Text Books:** NA	
11	**Reference Books:** 1. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling" by Ralph Kimball and Margy Ross. 2. Building the Data Warehouse" by W. H. Inmon, Richard D. Hackathorn, and Kent Graziano. 3. Data Warehouse Design: Modern Principles and Methodologies" by Matteo Golfarelli andStefano Rizzi. 4. Data Warehouse Design: Modern Principles and Best Practices" by Christopher Adamson andMichael Venerable. Data Warehousing in the Age of Big Data" by Krish Krishnan.	
	Lab Exercises – (30 Hours) Exercise 1: Downloading and Installing Talend Exercise 2: Launching Talend Exercise 3: Creating First Project Exercise 4: Working with a simple Job Exercise 5: CSV file as an Input Exercise 6: Edit Schema for CSV Input Exercise 7: Adding Logging Component Exercise 8: Propagating the Data and Running the CSV File as an Input Exercise 9: Excel as an Input and Output Exercise 10: Configure the components Exercise 11: Execute the Job Exercise 12: Sorting a File Exercise 13: Storing the Data & Adding Rules Exercise 14: Storing the Sorted Result in a File Exercise 15: Creating Metadata for Delimited File Exercise 16: Create a Job and Use the Created Metadata Exercise 17: Adding tLogRow and Run the Metadata Project Exercise 18: Creating a Metadata Database Connection Exercise 19: Reading a MySQL Table Using Metadata Exercise 20: Creating Project, Metadata and Job for tMap Exercise 21: Creating and Configuring tMap Exercise 22: Add tLogRow and Link Exercise 23: Run the tMap Job Exercise 24: Storing the tMap Output in a File	
12	**Internal Continuous Assessment: 40%**	**External, Semester End Examination 60% Individual Passing in Internal and External Examination**

13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem- III)
Title of Paper: Software Engineering

Sr. No.	Heading	Particulars
1	**Description the course :**	• Knowledge of basic SW engineering methods and practices, and their appropriate application. • Describe software engineering layered technology and Process frame work. • A general understanding of software process models such as the waterfall and evolutionary models.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	4 credits
5	**Hours Allotted :**	60 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Software Engineering (SE) comprises the core principles consistent in software construction and maintenance: fundamental software processes and life-cycles, mathematical foundations of software engineering, requirements analysis, software engineering methodologies and standard notations, principles of software architecture and re-use, software quality frameworks and validation, software development, and maintenance environments and tools.	
8	**Course Outcomes**: • Understanding of software requirements and the SRS documents. • Understanding of implementation issues such as modularity and coding standards. • Understanding on quality control and how to ensure good quality software.	
9	**Module I: (14 Hours)** Overview of System Analysis & Design, Business System Concept, System Development Life Cycle, Waterfall Model, Spiral Model, Feasibility Analysis, Technical Feasibility, Cost-Benefit Analysis, COCOMO model.	
	Module II (10 Hours) : System Design–Context diagram and, DFD, Problem Partitioning, Top-Down And Bottom-Up design; Decision tree, decision table and structured English; Functional v/s Object-Oriented approach.	
	Module III (14 Hours) : Coding & Documentation–Structured Programming, OO Programming, Information Hiding, Reuse, System Documentation. Testing–Levels of Testing, Integration Testing, Test case Specification, Reliability Assessment, Validation & Verification.	
	Module IV (10 Hours) : Software Project Management–Project Scheduling, Staffing, Software Configuration Management, Quality Assurance, Project Monitoring.	

	Module V (12 Hours) : Static and dynamic models, why modeling, UML diagrams: Class diagram, interaction diagram: collaboration diagram, sequence diagram, state chart diagram, activity diagram, implementation diagram.	
10	**Text Books: NA**	
11	**Reference Books:** 1. Pressman, Software Engineering: A practitioner's approach–(TMH) 2. Pankaj Jalote, Software Engineering-(Wiley-India) 3. N. S. Gill, Software Engineering–(Khanna Publishing House) 4. Rajib Mall, Software Engineering -(PHI)	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 10 Marks Assignments 10 Marks Mid Term Exam 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem - III)
Title of Paper: Internet Web and Programming

Sr. No.	Heading	Particulars
1	**Description the course :**	• Knowledge of basic SW engineering methods and practices, and their appropriate application. • Describe software engineering layered technology and Process frame work. • A general understanding of software process models such as the waterfall and evolutionary models.
2	**Vertical :**	VSC
3	**Type :**	Theory
4	**Credit**:	2 credits
5	**Hours Allotted :**	30 Hours
6	**Marks Allotted**:	50 Marks
7	**Course Objectives**: • On completion of this course, a student will be familiar with client server architecture and able to develop a web application using java technologies. • Students will gain the skills and project based experience needed for entry into web application and development careers.	
8	**Course Outcomes**: • Understand best technologies for solving web client/server problems • Analyze and design real time web applications • Use Java script for dynamic effects and to validate form input entry • Analyze to Use appropriate client-side or Server-side applications	
9	**Module I: (7 Hours)** Internet and WWW: What is Internet?, Introduction to internet and its applications, E- mail, telnet, FTP, e-commerce, video conferencing, e-business. Internet service providers, domain name server, and internet address World Wide Web (WWW): World Wide Web and its evolution, uniform resource locator (URL), browsers - internet explorer, netscape navigator, opera, firefox, chrome, mozilla. Search engine, web saver - apache, IIS, proxy server, HTTP protocol.	
	Module II (13 Hours) : HTML and Graphics : HTML Tag Reference, Global Attributes, Event Handlers, Document Structure Tags, Formatting Tags, Text Level formatting, Block Level formatting, List Tags, Hyperlink tags, Image and Image maps, Table tags, Form Tags, Frame Tags, Executable content tags. Imagemaps : What are Imagemaps? Client-side Imagemaps, Server-side Imagemaps, Using Server-side and Client-side Imagempas together, alternative text for Imagemaps, Tables : Introduction to HTML tables and their structure, The table tags, Alignment, Aligning Entire Table, Alignment within a row, Alignment within a cell, Attributes, Content Summary, Background color, Adding a Caption, Setting the width, Adding a border, Spacing within a cell, Spacing between the cells, spanning multiple rows or columns, Elements that can be placed in a table, Table Sections and column properties, Tables as a design tool Frames : Introduction to Frames, Applications, Frames document, The tag, Nesting tag, Placing content in frames with the tag, Targeting named frames, Creating floating frames, Using Hidden frames.	

	Forms : Creating Forms, The <FORM> tag,Named Input fields, The <INPUT> tag, Multiple lines text windows, Drop down and list boxes, Hidden, Text, Text Area, Password, File Upload, Button, Submit, Reset, Radio, Checkbox, Select, Option, Forms and Scripting, Action Buttons, Labelling input files, Grouping related fields, Disabled and read-only fields, Form field event handlers, Passing form data Style Sheets: What are style sheets?, Why are style sheets valuable? Different approaches to style sheets, Using Multiple approaches, Linking to style information in s separate file, Setting up style information, Using the <LINK> tag, embedded style information, Using <STYLE> tag, Inline style information	
	Module III (10 Hours) : Java Script: Introduction, Client-Side JavaScript, Server-Side JavaScript, JavaScript Objects, JavaScript Security, Operators: Assignment Operators, Comparison Operators, Arithmetic Operators, % (Modulus), ++ (Increment), -- (Decrement), -(Unary Negation), Logical Operators, Short-Circuit Evaluation, String Operators, Special Operators, ? (Conditional operator), ,(Comma operator), delete, new, this, void Statements : Break, comment, continue, delete, do … while, export, for, for…in, function, if…else, import, labelled, return, switch, var, while, with, Core JavaScript (Properties and Methods of Each) : Array, Boolean, Date, Function, Math, Number, Object, String, regExp Document and its associated objects : document, Link, Area, Anchor, Image, Applet, Layer Events and Event Handlers : General Information about Events, Defining Event Handlers, event, onAbort, onBlur, onChange, onClick, onDblClick, onDragDrop, onError, onFocus, onKeyDown, onKeyPress, onKeyUp, onLoad, onMouseDown, onMouseMove, onMouseOut, onMouseOver, onMouseUp, onMove, onReset, onResize, onSelect, onSubmit, onUnload	
10	**Text Books: NA**	
11	**Reference Books:** 1. Web Design The complete Reference, Thomas Powell, Tata McGrawHill 2. HTML and XHTML The complete Reference, Thomas Powell, Tata McGrawHill 3. JavaScript 2.0 : The Complete Reference, Second Edition by Thomas Powell and Fritz Schneider	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 5 Marks Presentation 5 Marks Assignments 5 Marks Mid Term Exam 5 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 10 Marks Question 2-7 each of 5 Marks (Attempt any 4)	

Sem. - IV

Semester IV	
Subject Name	**Credit**
Major (Mandatory)	
Descriptive Analytics	4
Machine Learning	4
Minor	
To be picked from University Basket	2
To be picked from University Basket	2
OE - Open Electives	
To be picked from University Basket	2
SEC - Skill Enhancement Course	
R-Programming	2
AEC - Ability Enhancement Course	
To be picked from University Basket	2
CEP - Continuing Education Program	
To be picked from University Basket	2
CC - Co-curricular Course	
To be picked from University Basket	2

Syllabus
B.Sc. (Data Science and Sports Studies)
(Sem - IV)

Semester IV

Sr. No.	Subjects	Subject Type	Credits	Internal	External	Total
1	Descriptive Analytics	Major	4	40	60	100
2	Machine Learning	Major	4	40	60	100
3	R-Programming	SEC	2	20	30	50
4	Pick from University Basket	Minor	2	20	30	50
5	Pick from University Basket	Minor	2	20	30	50
6	Pick from University Basket	OE	2	20	30	50
7	Pick from University Basket	AEC	2	20	30	50
8	Pick from University Basket	CEP	2	20	30	50
9	Pick from University Basket	CC	2	20	30	50

Syllabus
BSc. (Data Science and Sports Studies)
(Sem - IV)
Title of Paper: Descriptive Analytics

Sr. No.	Heading	Particulars
1	**Description the course :**	• Descriptive Analytics is a course designed to provide students with the knowledge and skills required to analyze and interpret historical data to gain insights into past events and trends. • The course emphasizes the use of various statistical and data visualization techniques to summarize and describe data effectively.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Understand the purpose and importance of descriptive analytics in data analysis. • Learn various statistical techniques for summarizing and interpreting data. • Develop skills in data visualization to effectively communicate findings.	
8	**Course Outcomes**: • Apply descriptive analytics techniques to real-world datasets. • Gain proficiency in using software tools for descriptive analytics.	
9	**Module I: (5 Hours)** Business intelligence, BI architecture & components, Functional areas of BI tool, Applications of BI tool, Enterprise analytics capabilities, Complex vs volume operations, The nature of industry: Online retailers, Examples of nature of industry, Complex vs volume operations, IT vs business and fact-based culture, A strong data infrastructure, Analytics workforce, Future of business intelligence.	
	Module II: (10 Hours) Decision support, Examples of decision support, Exploring and analyzing data, Data warehouse, Data warehouse properties, Types of data, Multidimensional data, Schemas, Star schemas, Snowflake schemas, Constellation schemas, Conceptual modeling of data warehouses, Data warehouse design process, Multi-tiered architecture, Cross tabulation of sales by item-name and color, Data cube, Hierarchies on dimensions, OLAP server architectures, Discovery-driven data cubes, Typical OLAP operations, OLTP vs. OLAP, Business query, Dashboards and scorecards development, Metadata model, Automated tasks and events, Mobile BI, Disconnected BI, Collaborative BI, Real-time monitoring Software Development Kit (SDK), Making BI easy to consume	

	Module III: (12 Hours) Business transformation projects, Where are we now? ASUG BI maturity model, Measuring BI success and value (McDonald 2004), BI effectiveness scorecard, BI value scorecard, Five key areas of strategy, Planning a BI Project, BI design and development, Engagement activities, Pre- engagement activities, Business environment, Engagement process, Project tasks Task 1: Knowledge capture goals - Discuss business objectives and prior learning, Interview key stakeholders Task 2: Consolidate findings Task 3: Map the customer situation - Current environment, Business/functional requirements sample diagram, Logical BI diagram Task 4: Methodology and approach Task 5: Standards and governance Task 6: Sections, milestones, and tasks Task 7: Proof of Concept (POC) Task 8: Table creation Task 9: OLAP creation Task 10: Three final deliverables Server administration data measurement, KDD: Knowledge Discovery in database, The Cross Industry Standard Process for Data Mining (CRISP- DM).
	Module IV: (12 Hours) Metrics, Senior managers should take care of metrics, Metrics drive behavior in several ways, Kaplan-Norton balanced scorecard, The Rayport- Jaworski performance dashboard and strategy framework, Introducing the R-J performance dashboard, Blueprint to the R-J performance dashboard, Building reports, Relational reporting style, Dimensional reporting style, Dashboard: Definition, Layers of information, Evolution of dashboards, Information discovery and delivery, Dashboard design principles, Chart overview, Singular components, Alerts overview, Example of a dashboard, Display media for dashboards, Sales simulation (what-if scenarios) What-if scenarios: Example 1: Calculation of a monthly rate for a home loan What-if scenarios: Example 2: Calculation of the average profit margin What-If scenarios: Input components What-if scenarios: Output components Setting up the business scenario for a sales simulation, List report, Crosstab report, Chart report, Map report, Data group and sort, Filters in report, Add calculations to report, Conditions in report, Adding summary values in report, Drilling in report: Drill up and down, Drilling in report: Drill through, Run report: On demand or schedule, Charts, Parts of a graph, Chart types, Chart type: Bar chart, Chart type: Line, Chart type: Pie, Chart type: Area, Chart type: Scatter
	Module V (6 Hours) : Data mining Data mining process Data mining Benefits Data mining disadvantages Data mining challenges Data mining examples Data mining methods
10	**Text Books: NA**
11	**Reference Books:** *1. Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.* *2. "Deep Learning with Python" by François Chollet.* *3. "Deep Learning for Computer Vision" by Adrian Rosebrock.* *4. "Deep Reinforcement Learning" by Pieter Abbeel and John Schulman. 5."Generative Deep Learning" by David Foster.* 5. *6."Deep Learning with PyTorch" by Eli Stevens, Luca Antiga, and Thomas Viehmann.*

	Lab Exercise: (30 Hours) Exercise 1: Power BI desktop Exercise 2: Power BI service Exercise 3: Architecture Exercise 4: Supported Data Sources Exercise 5: Data Modeling Exercise 6: Dashboard Options Exercise 7: Visualization Options 1 Exercise 8: Visualization Options 2 Exercise 9: Excel Integration Exercise 10: Sharing Power BI Dashboards1 Exercise 11: Sharing Power BI Dashboards 2 Exercise 12: DAX Basics 1 xercise 13: DAX Basics 2 xercise 14: DAX Basics 3 xercise 15: DAX Advanced 1 xercise 16: DAX Advanced 2 xercise 17: DAX Advanced 3 xercise 18: Starting with Shiny xercise 19: Build a Shiny app xercise 20: Customize reactions xercise 21: How to customize appearance xercise 22: Shiny Project xercise 23: Basic python operations xercise 24: Creating of Arrays xercise 25: Introduction to pandas package xercise 26: Introduction to NumPy xercise 27: Data merging xercise 28: Data Cleaning xercise 29: Data transformation xercise 30: Introduction to matplotlib xercise 31: Introduction to matplotlib - 2 xercise 32: Customizing Plots with Matplotlib xercise 33: Using Seaborn for data visualization xercise 34: Exploratory analysis – 1 xercise 35: Exploratory analysis – 2a xercise 36: Exploratory analysi2b xercise 37: Introduction to Web visualization using D3.js Exercise 38: Building Interactive Bar Charts	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

(Sem - IV)

Title of Paper: Machine Learning

Sr. No.	Heading	Particulars
1	**Description the course :**	• This course provides a comprehensive introduction to the theory and practice of machine learning. • Machine learning is a field of study that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. • This course covers a wide range of machine learning techniques and their applications, equipping students with the skills necessary to understand, implement, and evaluate machine learning models.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Understand the fundamental concepts and principles of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. • Familiarize with various types of machine learning algorithms, such as decision trees, support vector machines, neural networks, and ensemble methods. • Learn about data preprocessing techniques, including feature selection, feature scaling, and handling missing data. • Gain knowledge of different evaluation metrics and techniques to assess the performance of machine learning models. • Explore model selection and hyperparameter tuning techniques for optimizing machine learning algorithms.	
8	**Course Outcomes**: • Learn about different techniques for handling imbalanced datasets and dealing with class imbalance problems. • Gain practical experience in implementing machine learning algorithms using popular libraries and frameworks such as Scikit-learn and TensorFlow. • Apply machine learning techniques to real-world datasets and solve various prediction, classification, and clustering tasks. • Understand the ethical considerations and implications of machine learning, including fairness, transparency, and interpretability. • Explore advanced topics in machine learning, such as deep learning, natural language processing, and computer vision.	
9	**Module I: (7 Hours)** Introduction, Motivation for machine learning, Applications, Machine learning, Classification, Regression, The origin of machine learning, Time line of machine learning techniques, Uses and abuses of machine learning, How do machines learn, Abstraction and knowledge representation, Generalization, Assessing the success of learning, Steps to apply machine learning to data, Input data and ML algorithm, Machine learning methods, Unsupervised learning, Semi-supervised learning, Clustering, What are we looking for? Classification of machine learning algorithms, General ML architecture, Reinforcement learning, Supervised learning, Unsupervised learning, Semi-supervised learning, Regularization algorithms, Clustering algorithms, Deep learning algorithms, Ensemble learning, Matching data to an appropriate algorithm.	

	Module II (7 Hours) Introduction, Supervised learning, Regression, Regression examples, Regression models, Steps in regression analysis, Linear regression, Simple linear regression, Least squares estimation, Least squares regression-Line of best fit, Illustration, Direct regression method, Maximum likelihood estimation, Matrix approach, Regression assumptions and model properties, Coefficient of determination (R-squared), Example, Testing for significance, Testing hypothesis in simple linear regression, Illustration, Checking model adequacy, Over-fitting, Detecting over-fit models: Cross validation, Cross validation: The ideal procedure, Logistic regression.
	Module III (6 Hours) Introduction, Ordinary least squares estimation for multiple linear regression, Multiple linear regression model building, Partial correlation and regression model building, Multiple linear regression model, Interpretation of multiple linear regression coefficients-Partial regression coefficients, Standardized regression coefficients, Missing data, Validation of multiple regression model, Coefficient of multiple determination (R- Squared), Adjusted R-squared, Statistical significance of individual variables in multiple linear regression: t-Test.
	Module IV (10 Hours) Preamble: Machine learning, To classify faces and expressions, Introduction, ML classifier, Classification and general approach, Classification algorithms, Instance based learning, K-Nearest Neighbor, Decision trees, Attribute selection measure: Information gain, ID3 algorithm, Decision tree: weekend example, Converting a tree to rules, Bayesian algorithms, Ensemble, Stories of success, Why ensemble works? Ensemble of classifiers, Bagging, Boosting, Random forests, Neural networks, Activation functions, Feedforward neural network, Multi-layer perceptron, Backprop algorithm, Recurrent or feedback architecture, Perceptron rule, Gradient-descent (training examples, η), Multilayer networks and back propagation algorithm, Support vector machine, Classification model evaluation and selection, ROC curves, Cost Benefit Analysis (CBA).
	Module V (7 Hours) Clustering, Clustering algorithms, More common clustering situation, Statistics associated with cluster analysis, General applications of clustering, Clustering as a pre-processing tool, Hard vs. soft clustering, Similarity and dissimilarity between objects, Type of data in clustering analysis, Binary variables, Nominal variables, Ordinal variables, Major clustering approaches, Types of clusters, Cluster centroid and distances, Hierarchical clustering, Hierarchical Agglomerative Clustering (HAC), Hierarchical Agglomerative Clustering: Linkage method, Hierarchical Agglomerative Clustering: Variance and Centroid method, Cluster distance measures, Single link agglomerative clustering, Complete-link clustering, Average-link clustering, Other agglomerative clustering methods, Distance between two clusters, Hierarchical clustering: Time and Space requirements, K - means clustering, Importance of choosing initial centroids, The K-medoids clustering method, PAM (Partitioning Around Medoids), CLARA (Clustering Large Applications), CLARANS (Randomized CLARA), Density based clustering methods, DBSCAN: Density Based Spatial Clustering of Applications with Noise, When DBSCAN Does NOT Work Well, External criteria for clustering quality, Different aspects of cluster validation, Measures of cluster validity, Measuring cluster validity via correlation, Using similarity matrix for cluster validation, Internal measures: SSE, Framework for cluster validity, Internal measures: Cohesion and Separation, Internal measures: Silhouette coefficient.
	Module VI (8 Hours) Information retrieval: introduction, Information retrieval process, Information retrieval architecture, how do we represent document? Information retrieval models, Similarity metric, Term weighting, Retrieval in vector space model, Constructing inverted index (word counting), Stop words removal, Stemming, Text document clustering, Agglomerative vs. divisive, Impact of cluster distance measure, Buckshot clustering, Issues related to cosine similarity, Validity of document clusters, Text datasets, Experimental evaluation.
10	**Text Books: NA**

11	**Reference Books:** 1. The Hundred-Page Machine Learning Book by Andriy Burkov. 2. Machine Learning For Absolute Beginners by Oliver Theobald. 3. Programming Collective Intelligence by Toby Segaran. 4. Machine Learning for Humans by Vishal Maini and Samer Sabri.	
	Lab Exercises –(30 hours) Exercise 1: Linear Regression Exercise 2: Best Fit for LR Exercise 3: Logistic Regression Exercise 4: Logistic Regression - New Product Purchase Dataset Exercise 5: Multiple Linear Regression Exercise 6: Multiple Linear Regression Using Random Data Exercise 7: KNN Accuracy Prediction Exercise 8: KNN algorithm implementation on Breast Cancer Data Exercise 9: KNN algorithm implementation with random value and its power calculation Exercise 10: Implement a Decision Tree using Balance Scale Dataset Exercise 11: Implement Naïve Bayes Classification Using iris data Exercise 12: Implement Support vector machine using random dataset Exercise 13: Implement Principal component analysis with Wine dataset Exercise 14: Implement Bagging using Sonar Dataset Exercise 15: Implement Boosting using Mushroom Dataset and AdaBoost Classifier Exercise 16: DBSCAN with credit card Dataset	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem - IV)
Title of Paper: R Programming

Sr. No.	Heading	Particulars
1	**Description the course :**	• In this course students will learn how to program in R and how to use R for effective data analysis. • You will learn how to install and configure software necessary for a statisticalprogramming environment,

		discuss generic programming language concepts as they are implemented in a high-level statistical language.
2	**Vertical :**	VSEC
3	**Type :**	Theory
4	**Credit**:	2 credits
5	**Hours Allotted :**	30 Hours
6	**Marks Allotted**:	50 Marks
7	**Course Objectives**: • To learn basic concept of R programming • To learn basic and advanced data structure in R Programming • To learn and implement advanced concept of data visualization techniques using R	
8	**Course Outcomes**: • The course covers practical issues in statistical computing which includes programming in R, reading data into R, accessing R packages, writing R functions, debugging, and organizing and commenting R code.	
9	**Module I: (4 Hours)** Introduction to R – Help Functions in R – Vectors – Vectorized Operations –Functions in R – Packages in R.	
	Module II (8 Hours) : Matrices, Arrays and Lists Matrix Operations – Adding and Deleting Rows and Columns – Higher Dimensional Arrays – Lists – General List Operations – Accessing List Components and Values – Applying functions to Lists	
	Module III (6 Hours) : Data Frames Creating Data Frames – Matrix-like Operations on a Data Frame – Merging Data Frames – Applying functions to Data Frames – Factors and Tables – Common Functions used with Factors – Working withTables.	
	Module IV (6 Hours) : OOP S3 Classes – S4 Classes – Managing the Objects – Input / Output – Accessing Keyboard and Monitor – Reading and Writing Files – accessing the Internet – String Manipulation.	
	Module V (4 Hours) : Data Visualization Introduction to GGPlot2 – Factors – Aesthetics – Plotting with Layers – Overriding Aesthetics – Mapping vs Setting – Histograms – Density Charts – Statistical Transformation – Facets – Coordinates – Themes.	
10	**Text Books: NA**	
11	**Reference Books:** 1. Norman Matloff , "The Art of R Programming: A Tour of Statistical Software Design", No Starch Press, 2011. 2. Jared P. Lander, "R for Everyone: Advanced Analytics and Graphics", Addison-Wesley Data & Analytics Series, 2013. 3. Mark Gardener, " Beginning R – The Statistical Programming Language", Wiley, 2013 4. Robert Knell, "Introductory R: A Beginner's Guide to Data Visualisation, Statistical Analysis and Programming in R", Amazon Digital South Asia Services Inc, 2013.	

12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignment 05 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 10 Marks Question 2-7 each of 05 Marks (Attempt any 4)	

Sem. - V

Semester V	
Subject Name	**Credit**
Major (Mandatory)	
Predictive Analytics	4

Big Data Analytics	4
Project 1	2
Minor	
To be picked from University Basket	2
To be picked from University Basket	2
OE - Open Electives	
To be picked from University Basket	2
SEC - Skill Enhancement Course	
To be picked from University Basket	2
AEC - Ability Enhancement Course	
To be picked from University Basket	2
FP - Field Project	
To be picked from University Basket	2

Syllabus
B.Sc. (Data Science and Sports Studies)
(Sem.- IV)

Semester IV

Sr. No.	Subjects	Subject Type	Credits	Internal	External	Total
1	Predictive Analytics	Major	4	40	60	100
2	Big Data Analytics	Major	4	40	60	100
3	Project 1	Major	2	20	30	50
4	Distributed Systems	VSC	2	20	30	50
5	Pick from University Basket	Minor	4	40	60	100
6	Pick from University Basket	FP	2	20	30	50
7	Pick from University Basket	CEP	2	20	30	50

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Predictive Analytics

Sr. No.	Heading	Particulars
1	**Description the course :**	• This course provides a comprehensive introduction to data mining techniques and their applications. • Students will learn the process of extracting valuable insights from diverse datasets, including data preparation, model development, evaluation, and deployment. • The course covers a wide range of topics, from linear regression to neural networks and web mining, equipping students with practical skills for analyzing and interpreting data. • Through a mix of theoretical concepts and hands-on exercises, students will gain the necessary skills to apply data mining techniques effectively.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Understand the fundamentals of data mining and its significance. • Gain practical experience in data preparation and exploration techniques. • Develop proficiency in building and evaluating predictive models. • Explore a variety of data mining techniques, from regression to neural networks.	
8	**Course Outcomes**: • Learn about web mining and its applications in analyzing web usage data. • Apply data mining skills to real-world scenarios and decision-making. • Understand the challenges and considerations in predictive analytics. • Explore the growth and applications of machine learning.	
9	**Unit I: (12 Hours)** Introduction to Data Mining - Significance of data mining in extracting insights. - Evolution of data mining techniques. Understanding Data Mining - Defining data mining and its purpose. -Need for data mining tools. Data Mining Process and KDD Model - Process of data mining and Knowledge Discovery in Databases (KDD) model. - Research challenges in the KDD process. Types of Data for Mining - Different kinds of data suitable for data mining. - Scenario: Need for databases in data mining. Types of Data Mining Tasks and CRISP-DM	

	- Overview of data mining tasks: descriptive, predictive, etc. - Introduction to CRISP-DM methodology. Components and Techniques in Data Mining - Components of data mining methods. - Various data mining operons and techniques. Applications of Data Mining and Predictive Analytics - Practical applications of data mining across domains. - Predictive analytics and its usage.techniques, Practical examples/case studies.
	Unit II (12 Hours) Data Preparation and Quality - Importance of data preparation in data mining. - Data quality and its significance. • Data Collection Methods and Sampling - Techniques for data collection and sampling. - Ensuring data understanding and quality. • Data Exploration and Visualization - Categories of data visualization for exploration. - Verifying data quality through outlier detection. • Data Cleaning and Transformation - Cleaning data: Acquisition and unified date format. - Handling missing data and categorical variables. • Univariate Data Analysis and Statistics - Performing univariate data analysis. - Analyzing continuous variable distribution, standard deviation, and percentiles. • Analysis of Categorical Data and Discretization - Analyzing observed vs. expected distribution. - Discretization techniques without using classes.
	Unit III (9 Hours) Data Partitioning and Model Selection - Techniques for data partitioning and model selection. • Linear Regression and Decision Trees - Building a Simple Linear Regression (SLR) model. - Introduction to decision trees and their algorithm. • Clustering, Link Analysis, and Classification - Understanding clustering and link analysis. - Introduction to classification techniques. • Decision Trees and Information Theoretic Approach - Decision tree algorithm and key requirements. - Information theoretic approach for decision trees. • Classification Techniques and SVM - Techniques for classification, including Support Vector Machine (SVM)
	Unit IV (12 Hours) Model Evaluation Methods - Techniques for model evaluation and comparison. • Model Evaluation Metrics - Using evaluation charts for model comparison. - Confusion matrix and measures of accuracy. • Liti Charts, Gain Charts, and ROC - Analyzing liŌ charts, gain charts, and Receiver Operating Characteristic (ROC) curve. • Regression Evaluation and Meta-level Modeling - Evaluating regression models and meta-level modeling. • Model Deployment and Performance Assessment - Deploying models and assessing their performance. • Model Selection and Assessment - Assessing bias, variance, and model complexity. - Guidelines for model selection using AIC, BIC, and cross- validation. • Model Validation and Data Collection Strategies - Validating models and potential issues in model selection. - Strategies for data collection in regression model building.
10	**Text Books:**
11	**Reference Books:** 5. 1.M. Kantardzic, "Data mining: Concepts, models, methods and algorithms, John Wiley &Sons Inc. 6. Zantinge , "Data Mining", Pearson Education Asia 7. 3. J. Han, M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann
	Lab Exercises – (30 Hours) 1. Design and Create Cube by identifying measures and dimensions for Star Schema, Snowflake 2. Design and Create Cube by identifying measures and dimensions for Design storage for cube using storage

	3. Process Cube and Browse Cube Data 1. By replacing a dimension in the grid, filtering and drilldown using cube browser 2. Browse dimension data and view dimension members, member properties, member property values 4. Create calculated member using arithmetic operators and member property of dimension Member 5 Design and Create data mining models using Analysis Service of SQL server 2005. 6 Design and Build targeted mailing data mining model using analysis service of SQL server 2005 and compare their predictive capabilities using the Mining Accuracy Chart View and Create predictions using Prediction Query Builder. 7 Perform various steps of Preprocessing on the given relational database / warehouse. 8 To implement Data Mining Extensions (DMX) language and MDX query language 9 Perform various steps of Preprocessing using WEKA software	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Machine Learning

Sr. No.	Heading	Particulars
1	**Description the course :**	• This course provides a comprehensive introduction to the theory and practice of machine learning. • Machine learning is a field of study that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. • This course covers a wide range of machine learning techniques and their applications, equipping students with the skills necessary to understand, implement, and evaluate machine learning models.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Understand the fundamental concepts and principles of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. • Familiarize with various types of machine learning algorithms, such as decision trees, support vector machines, neural networks, and ensemble methods. • Learn about data preprocessing techniques, including feature selection, feature scaling, and handling missing data. • Gain knowledge of different evaluation metrics and techniques to assess the performance of machine learning models. • Explore model selection and hyperparameter tuning techniques for optimizing machine learning algorithms.	
8	**Course Outcomes**: • Learn about different techniques for handling imbalanced datasets and dealing with class imbalance problems. • Gain practical experience in implementing machine learning algorithms using popular libraries and frameworks such as Scikit-learn and TensorFlow. • Apply machine learning techniques to real-world datasets and solve various prediction, classification, and clustering tasks. • Understand the ethical considerations and implications of machine learning, including fairness, transparency, and interpretability. • Explore advanced topics in machine learning, such as deep learning, natural language processing, and computer vision.	

9	**Module I: (7 Hours)** Introduction, Motivation for machine learning, Applications, Machine learning, Classification, Regression, The origin of machine learning, Time line of machine learning techniques, Uses and abuses of machine learning, How do machines learn, Abstraction and knowledge representation, Generalization, Assessing the success of learning, Steps to apply machine learning to data, Input data and ML algorithm, Machine learning methods, Unsupervised learning, Semi-supervised learning, Clustering, What are we looking for? Classification of machine learning algorithms, General ML architecture, Reinforcement learning, Supervised learning, Unsupervised learning, Semi-supervised learning, Regularization algorithms, Clustering algorithms, Deep learning algorithms, Ensemble learning, Matching data to an appropriate algorithm.
	Module II (7 Hours) Introduction, Supervised learning, Regression, Regression examples, Regression models, Steps in regression analysis, Linear regression, Simple linear regression, Least squares estimation, Least squares regression-Line of best fit, Illustration, Direct regression method, Maximum likelihood estimation, Matrix approach, Regression assumptions and model properties, Coefficient of determination (R-squared), Example, Testing for significance, Testing hypothesis in simple linear regression, Illustration, Checking model adequacy, Over-fitting, Detecting over-fit models: Cross validation, Cross validation: The ideal procedure, Logistic regression.
	Module III (6 Hours) Introduction, Ordinary least squares estimation for multiple linear regression, Multiple linear regression model building, Partial correlation and regression model building, Multiple linear regression model, Interpretation of multiple linear regression coefficients-Partial regression coefficients, Standardized regression coefficients, Missing data, Validation of multiple regression model, Coefficient of multiple determination (R- Squared), Adjusted R-squared, Statistical significance of individual variables in multiple linear regression: t-Test.
	Module IV (10 Hours) Preamble: Machine learning, To classify faces and expressions, Introduction, ML classifier, Classification and general approach, Classification algorithms, Instance based learning, K-Nearest Neighbor, Decision trees, Attribute selection measure: Information gain, ID3 algorithm, Decision tree: weekend example, Converting a tree to rules, Bayesian algorithms, Ensemble, Stories of success, Why ensemble works? Ensemble of classifiers, Bagging, Boosting, Random forests, Neural networks, Activation functions, Feedforward neural network, Multi-layer perceptron, Backprop algorithm, Recurrent or feedback architecture, Perceptron rule, Gradient-descent (training examples, η), Multilayer networks and back propagation algorithm, Support vector machine, Classification model evaluation and selection, ROC curves, Cost Benefit Analysis (CBA).
	Module V (7 Hours) Clustering, Clustering algorithms, More common clustering situation, Statistics associated with cluster analysis, General applications of clustering, Clustering as a pre-processing tool, Hard vs. soft clustering, Similarity and dissimilarity between objects, Type of data in clustering analysis, Binary variables, Nominal variables, Ordinal variables, Major clustering approaches, Types of clusters, Cluster centroid and distances, Hierarchical clustering, Hierarchical Agglomerative Clustering (HAC), Hierarchical Agglomerative Clustering: Linkage method, Hierarchical Agglomerative Clustering: Variance and Centroid method, Cluster distance measures, Single link agglomerative clustering, Complete-link clustering, Average-link clustering, Other agglomerative clustering methods, Distance between two clusters, Hierarchical clustering: Time and Space requirements, K - means clustering, Importance of choosing initial centroids, The K-medoids clustering method, PAM (Partitioning Around Medoids), CLARA (Clustering Large Applications), CLARANS (Randomized CLARA), Density based clustering methods, DBSCAN: Density Based Spatial Clustering of Applications with Noise, When DBSCAN Does NOT Work Well, External criteria for clustering quality, Different aspects of cluster validation, Measures of cluster

	validity, Measuring cluster validity via correlation, Using similarity matrix for cluster validation, Internal measures: SSE, Framework for cluster validity, Internal measures: Cohesion and Separation, Internal measures: Silhouette coefficient.	
	Module VI (8 Hours) Information retrieval: introduction, Information retrieval process, Information retrieval architecture, how do we represent document? Information retrieval models, Similarity metric, Term weighting, Retrieval in vector space model, Constructing inverted index (word counting), Stop words removal, Stemming, Text document clustering, Agglomerative vs. divisive, Impact of cluster distance measure, Buckshot clustering, Issues related to cosine similarity, Validity of document clusters, Text datasets, Experimental evaluation.	
10	**Text Books: NA**	
11	**Reference Books:** 5. The Hundred-Page Machine Learning Book by Andriy Burkov. 6. Machine Learning For Absolute Beginners by Oliver Theobald. 7. Programming Collective Intelligence by Toby Segaran. 8. Machine Learning for Humans by Vishal Maini and Samer Sabri.	
	Lab Exercises –(30 hours) Exercise 1: Linear Regression Exercise 2: Best Fit for LR Exercise 3: Logistic Regression Exercise 4: Logistic Regression - New Product Purchase Dataset Exercise 5: Multiple Linear Regression Exercise 6: Multiple Linear Regression Using Random Data Exercise 7: KNN Accuracy Prediction Exercise 8: KNN algorithm implementation on Breast Cancer Data Exercise 9: KNN algorithm implementation with random value and its power calculation Exercise 10: Implement a Decision Tree using Balance Scale Dataset Exercise 11: Implement Naïve Bayes Classification Using iris data Exercise 12: Implement Support vector machine using random dataset Exercise 13: Implement Principal component analysis with Wine dataset Exercise 14: Implement Bagging using Sonar Dataset Exercise 15: Implement Boosting using Mushroom Dataset and AdaBoost Classifier Exercise 16: DBSCAN with credit card Dataset	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Big Data Analytics

Sr. No.	Heading	Particulars
1	**Description the course :**	• The subjects the fundamentals and advanced topics of Big Data Analytics along with use cases.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	4 credits
5	**Hours Allotted :**	60 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Demonstrates the basic operation in bigdata. • Awareness of Architectures use in bigdata. • Study about Big Data processing language Hadoop.	
8	**Course Outcomes**: • Describes about the modern databases. • Case studies in big data are analyzed.	
9	**Unit I: (20 Hours)** Fundamentals of Big-data analytics, Overview & analytics life cycle, Need, Structured and multi-structured data analysis, Big-data analytics major components, Analytical models and approaches, Relational and non- relational Databases, Application areas, Design and analysis of Analytics model-Analytics design steps, Understanding different data processing models, Statistical models, Predictive models, Descriptive models, Introduction to Data analysis using R- Basics of R Language, Statistical models in R, Statistical techniques applied using R, Graphical analysis techniques, Practical examples/case studies.	
	Unit II (25 Hours) : Analytics Using MapReduce, MapReduce design patterns, Clustered Hadoop environment, Advanced HDFS, Graph Algorithms, Searching and Indexing approaches, MapReduce Applications, Introduction to Pig and HIVE- Programming Pig: Engine for executing data flows in parallel on Hadoop, Programming with Hive: Data warehouse system for Hadoop	
	Unit III (15 Hours) : Case Studies -Real time stream processing using MapReduce and R and Text Analytics based on different documents for identifying interesting patterns and correlations	
10	**Text Books: NA**	
11	**Reference Books:** 5. M. Kantardzic, "Data mining: Concepts, models, methods and algorithms, John Wiley &Sons Inc. 6. Adriaans, Dolf Zantinge , "Data Mining", Pearson Education Asia 7. J. Han, M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann	

12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 10 Marks Assignment 10 Marks Oral Test 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Project 1

Sr. No.	Heading	Particulars
1	**Description the course :**	• The Project Implementation course in Analytics builds upon the knowledge and skills acquired in previous semesters, focusing on the practical implementation of data science projects. • Students will work on real-world projects, applying data science methodologies and techniques to solve complex problems. • They will gain hands-on experience in data collection, preprocessing, feature engineering, and model development.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	2 credits
5	**Hours Allotted :**	30 Hours
6	**Marks Allotted**:	50 Marks
7	**Course Objectives**: • To enable students to apply the knowledge and skills acquired during the B.Sc Data Science program in the implementation of a data science project. • To provide students with hands-on experience in implementing data science methodologies and techniques to solve real-world problems.	
8	**Course Outcomes**: • To develop students' project management skills by planning, executing, and monitoring a data science project. • To enhance students' collaboration and teamwork abilities by working in groups to implement a data science project. • To foster students' critical thinking and problem-solving skills by addressing challenges and making informed decisions during the project implementation phase.	
9	**Unit I: (30 Hours)** Students can undertake industrial project.	
10	**Text Books: NA**	
11	**Reference Books:**	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks	

	Presentation 10 Marks Assignment 10 Marks Oral Test 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Distributed Systems

Sr. No.	Heading	Particulars
1	**Description the course :**	• The objective of this courses to provide the students a knowledge of c programmingenvironment to create, implement and debug programs.
2	**Vertical :**	VSC
3	**Type :**	Theory
4	**Credit**:	2 credits
5	**Hours Allotted :**	30 Hours
6	**Marks Allotted**:	50 Marks
7	**Course Objectives**: • Students will identify the core concepts of distributed systems: the way in which several machines orchestrate to correctly solve problems in an efficient, reliable and scalable way.	
8	**Course Outcomes**: • Learn the principles, architectures, and algorithms used in distributed systems, including communication, synchronization, resource sharing, and fault tolerance. • Explore key concepts such as concurrency, scheduling, and abstraction.	
9	**Unit I: (10 Hours)** Introduction and types of distributed systems – Architecture of DS - Overview of processes - Taxonomy of Distributed Systems - scalable performance - load balancing and availability. Models of computation - shared memory and message passing system - synchronous and asynchronous systems. Communication in Distributed Systems - Remote Procedure Calls and Message Oriented Communications and implementation - High-level communication and publish-subscribe in Map reduce.	
	Unit II (10 Hours) : Logical time and event ordering - Global state and snapshot algorithms - distributed snapshots in VMs - clock synchronization - Distributed mutual exclusion - Group based Mutual Exclusion - leader election - deadlock detection - termination detection - Distributed Databases - implementations over a simple distributed system and case studies of distributed databases and systems - Distributed file systems: scalable performance, load balancing, and availability. Examples from Dropbox, Google FS (GFS)/ Hadoop Distributed FS (HDFS), Bigtable/HBase MapReduce, RD	
	Unit III (10 Hours) : Consistency control: Data Centric Consistency - Client Centric Consistency - Replica Management - Consistency Protocols. Fault tolerance and recovery: basic concepts - fault models - agreement problems and its applications - commit protocols - voting protocols - check pointing and recovery. Case Studies from Apache Spark, Google Spanner, Amazon Aurora, Block Chain Systems etc.	
10	**Text Books: NA**	

11	**Reference Books:** 1. M. Van Steen, A.S. Tanenbaum, Distributed Systems, Third Edition, CreateSpace Independent Publishing Platform, 2017. 2. Ajay D. Kshemkalyani and Mukesh Singhal, Distributed Computing: Principles, Algorithms and Systems, Cambridge University Press, 2011. 3. Garg VK. Elements of distributed computing. John Wiley & Sons, 2002. 4. George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair, Distributed Systems: Concepts and Design, Fifth Edition, Pearson Education, 2017	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Quizzes 05 Marks Assignment 05 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 10 Marks Question 2-7 each of 5 Marks (Attempt any 4)	

Sem. - VI

Semester V	
Subject Name	**Credit**
Major (Mandatory)	
Social and Web Analytics	4
Analytics for Industries	4
Project 2	2
Minor	
To be picked from University Basket	2
To be picked from University Basket	2
OE - Open Electives	
To be picked from University Basket	2
SEC - Skill Enhancement Course	
To be picked from University Basket	2
AEC - Ability Enhancement Course	
To be picked from University Basket	2
FP - Field Project	
To be picked from University Basket	2

Syllabus
B.Sc. (Data Science and Sports Studies)
(Sem.- VI)

Semester VI

Sr. No.	Subjects	Subject Type	Credits	Internal	External	Total
1	Social and Web Analytics	Major	4	40	60	100
2	Analytics for Industries	Major	4	40	60	100
3	Project 2	Major	2	20	30	50
4	Pick from University Basket	VSC	2	20	30	50
5	Pick from University Basket	Minor	4	40	60	100
6	Pick from University Basket	FP	2	20	30	50
8	Pick from University Basket	CEP	2	20	30	50

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- VI)
Title of Paper: Social and Web Analytics

Sr. No.	Heading	Particulars
1	**Description the course :**	• Social and Web Analytics is a course that focuses on understanding and analyzing data from online sources, including social media platforms and websites. • The course explores various techniques and tools to extract insights from social and web data, understand user behavior, and make informed decisions based on the analysis.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	4 credits
5	**Hours Allotted :**	60 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Understand the importance and impact of social and web analytics in today's digital landscape. • Learn techniques for collecting, processing, and analyzing social and web data.	
8	**Course Outcomes**: • Develop skills in social media monitoring, sentiment analysis, and web traffic analysis. • Apply social and web analytics techniques to real-world datasets. • Gain proficiency in using relevant tools and platforms for social and web analytics.	
9	**Unit I: (20 Hours)** Social media, Importance of social media, Social media popularity, Benefits of social media, Tools available, Need of using analytics, Social analytics vs web analytics, Web analytics, Terms used by web analytics tools, Types of web analytics, Social and web analytics technical requirements, Social analytics main basic activities, Social media environment, Impact of social media on business, Leverage social media for better services, Current analytics platforms, Open-source vs licensed platforms, Google analytics, IBM social media analytics, Other social and web analytics tools, Choosing right specifications and optimal solution, Measuring variations in user behavior, The diversity of user activities, The origin of the user activity distribution, Long tails everywhere: The 80/20 rule (p/q Rule), User activities on twitter.	
	Unit II (13 Hours) : Data collection strategy, Content sharing, Challenges in the social media data collection, Characteristics of people-centric approach, Social graph, Influencers, Community managers, Organizing for social media, Factors that help in applying the model to an organization, Measures to organize the social media for success, Choosing focused data sources and social networks, Collecting and understanding social media data, Social data collection strategy, Facebook social APIs, Best practices for collecting data from social networks, Leverage qualitative data: What, why and how much, How to conduct a lab usability testing, Usability alternatives, Recruiting user research participants with social media, Outsourced online usability, Web- enabled emerging user research options, Online surveys, Types of surveys,	

	Eight tips to choose online survey providers, Types and Properties of Social Networks, Directed Versus Undirected Graphs, Node and Edge Properties, Visualization Network, The discipline of social media analytics, Collecting data, Performing analysis, Reporting results, Aligning social objectives with business goals, Social media objectives, Developing a waterfall strategy, A SMART methodology, Identify common social business objectives, Developing KPIs, Creating KPI's, Developing KPI's, Basic KPIs, Standard vs critical metrics , Critical web metrics: Bounce rate, Critical web metrics: Exit rate, Critical web metrics: Conversion rate, Critical web metrics: Engagement, Strategically aligned KPIS and tactics to find best web and social metrics, Specific KPIs, Tactics to find the best web and social media metrics, Business goals, Visualizing the social analytics framework, Build scorecards and dashboards to track KPIs, Measuring macro and micro conversions, Quantify economic value, Measuring success for non-ecommerce and B2B websites, Measuring B2B websites.	
	Unit III (12 Hours) : Managing web and social analytics, Social media analytics, Dashboard, Dashboard: Drill down capabilities, Dashboards in IBM Cognos Consumer Insight (CCI), Types of charts, Drill down to view more details, Sharing insights with dashboards, Relationships, Viewing relationships, Sentiments, Sentiment analysis, Sentiment terms, Evolving topics, Evolving topics: IBM social media analytics, Reports, Social media impact, Segmentation, Discovery, Content creation and tracking, Social media content creation, Issues in content creation, Competitive intelligence analysis, CI data sources and types, Competitor intelligence studies, Website traffic analysis, Analyzing referrals and destinations, Top keywords performance trend, Audience identification and segment analysis, Segmentation analysis, Social media optimization and benefits, Social media enablement audit, Understand signals and potential, Improving social signals, Focus on text and unstructured data, The basics of natural language processing, Modelling: Unsupervised vs supervised.	
	Unit IV (15 Hours) : Future of social media analytics and monitoring, , Large volumes of social media, Future of social analytics, Future social media analytics and monitoring, Mashing up data from disparate sources, Integrating customer profile data, True profile, Colliding data sets for big bang ideas, Purchase funnel, Social interactions and web visits, Server response time and help desk tickets, Social media integration and benefits, Integrate solution to share outcome with others, Strategy planning, Benefits of social media integration, Integrating social media into the organization, Integration mode, Social media integration, Goals of integrating social media, Social media integration, Social media services online, Search engines, Content engagement, Interactions: Real world & people, Learning and mapping, Under and over fitting, Regularizing in matrix factorization, Exploratory analysis, Prediction and recommendation.	
10	**Text Books: NA**	
11	**Reference Books:** *6. The Hundred-Page Machine Learning Book by Andriy Burkov* *7. Machine Learning For Absolute Beginners by Oliver Theobald* *8. Programming Collective Intelligence by Toby Segaran* 9. *Machine Learning for Humans by Vishal Maini and Samer Sabri*	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentations 10 Marks Assignment 10 Marks Oral Exam 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 10 Marks Question 2-7 each of 5 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- VI)
Title of Paper: Analytics for Industries

Sr. No.	Heading	Particulars
1	**Description the course :**	• This course provides a comprehensive exploration of Business Analytics (BA) across diverse industries, including Banking, Insurance, Telecom, and Healthcare. Through a combination of theoretical insights and practical examples, students will gain a deep understanding of the role of Analytics in addressing industry specific challenges, Optimizing decision-making processes, and fostering business growth. • The course delves into various types of Analytics, the significance of a fact-based decision-making culture, and the technologies that drive Analytics - driven transformations in industries.
2	**Vertical :**	Major
3	**Type :**	Theory & Practical
4	**Credit**:	4 credits
5	**Hours Allotted :**	75 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Comprehend the fundamental concepts of Business Analytics and its applications in different industry contexts. • Differentiate between Descriptive, Predictive, Prescriptive, and Cognitive analytics and their significance. • Analyze challenges posed by data volume, variety, and speed in industries, and propose effective solutions.	
8	**Course Outcomes**: • Appraise the importance of enterprise analytics capabilities and data infrastructure for decision-making. • Apply the IBM Business Analytics Maturity Model to assess an organization's analytics readiness. • Evaluate the role of optimization in enhancing business outcomes within industries. • Understand the IBM Technology Porŋolio for Business Analytics and its relevance to specific industries.	
9	**Unit I: (10 Hours)** Overview of Business Analytics - Introduction to business Analytics and its significance. - Types of Analytics: Descriptive, Predictive, Prescriptive, And Cognitive. • Business Analytics Capabilities and Technologies - Enterprise Analytics capabilities and technologies. - Predictive Analytics, prescriptive Analytics, cognitive Analytics. • Business Analytics Challenges and Approaches - Addressing challenges: Volume, Variety, Speed of Data. - Building a fact-based decision-making culture. - Importance of data infrastructure and analytical tools. • Business Analytics Maturity Model and Optimization - IBM Business	

	Analytics Maturity Model. - Key concepts and importance of optimization. • IBM Technology Porɳolio for Business Analytics - IBM technologies for business Analytics and optimization.
	Unit II (13 Hours) Banking Industry Overview - Challenges and priorities in the banking industry. • Business Analytics Capabilities for Banking - Enhancing competitiveness using business Analytics . - Examples of actionable insights in banking. • Business Analytics Applications in Banking - Operational efficiency, integrated risk management, customer care. - Payments and securities: Real-world examples. • Business Analytics Maturity Model for Banking - Applying the IBM Business Analytics Maturity Model to banking. • IBM Technology Porɳolio for Banking - IBM solutions and technologies for business Analytics in banking.
	Unit III (12 Hours) Changes in the Insurance Industry - Evolving challenges and opportunities in insurance. • Business Analytics for Customer Focus - Creating a customer-focused enterprise. - Leveraging customer insight for retention and growth. • Business Analytics for Risk Management - Integrated risk management using Analytics. • Business Analytics for Sales and Claims - Optimizing sales performance and claims process. • IBM Technology Porɳolio for Insurance - IBM solutions and technologies for business Analytics in insurance.
	Unit IV (10 Hours) Challenges in the Telecom Industry - Addressing challenges in a connected world. • Smarter Analytics for Telecom - Leveraging Analytics for competitive advantage. • Business Analytics Applications in Telecom - Customer Analytics, marketing optimization, and social media. - Network Analytics for insights and performance. • IBM Technology Porɳolio for Telecom - IBM solutions and technologies for business Analytics in telecom.
10	**Text Books: NA**
11	**Reference Books:** 9. "Analytics in a Big Data World: The Essential Guide to Data Science and its Applications " by Bart Baesens 10. "Big Data Analytics in Banking and Finance" by Gloria Phillips-Wren and Stephanie Teufel2. "Business Analytics for Managers" by Gert H. N. Laursen and Jesper Thorlund 11. "Predictive Analytics : The Power to Predict Who Will Click, Buy, Lie, or Die" by Eric Siegel 12. "Big Data Analytics : Turning Big Data into Big Money" by Frank J. Ohlhorst 13. "Predictive Analytics for Insurance" by Eric Siegel 14. "Business Analytics for Telecoms: Building a BeΣer Customer Experience" by Tony Costa and Yves 15. de Montcheuil 16. "Healthcare Analytics for Quality and Performance Improvement" by Trevor L. Strome 17. "Healthcare Analytics : From Data to Knowledge to Healthcare Improvement" by Hui Yang and Walter W. Piegorsch
	Lab Exercises – (30 hours) Exercise 1. Twitter Data Analysis R Exercise 2. What's Happening with 140 Characters R (SHOW WHEN THE TWEETS ARE PUBLISHED) Exercise 3. Trend analysis R (SHOW WHEN THE TWEETS ARE PUBLISHED) Exercise 4. Sentiment analysis R Exercise 5. Follower graph analysis (SHOW THE ACCOUNTS FROM WHICH MOST RETWEETS ORIGINATE)

	Exercise 6. Analyzing Social Networks and Brand Engagements R (SHOW THE MOST FREQUENTLY USED HASHTAGS) Exercise 7. Facebook analyses using R (basis descriptive statistics using own profile) Exercise 8. Facebook analyses using R (Finding location and other details of friend's profile) Exercise 9. Analyzing English Football Club's brand page engagements Exercise 10. Analyzing English Football Club's brand page engagements (Visualizing post counts per page) Exercise 11. Analyzing English Football Club's brand page engagements (Visualizing post counts by post type per page) Exercise 12. Analyzing English Football Club's brand page engagements (Visualizing average likes by post type per page) Exercise 13. Analyzing English Football Club's brand page engagements (Visualizing average shares by post type per page) Exercise 14. Analyzing English Football Club's brand page engagements (Visualizing page engagement over time) Exercise 15. Analyzing English Football Club's brand page engagements (Visualizing user engagement with page over time) Exercise 16. Analyzing English Football Club's brand page engagements (Trending posts by user likes per page) Exercise 17. Google analytics setup Exercise 18. Google Analytics - Real-time Operations Google Analytics Exercise 19. Google Analytics - Audience Analysis Google Analytics Exercise 20. Google Analytics - Acquisition analysis Google Analytics Exercise 21. LinkedIn data analysis using R	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 05 Marks Assignments 05 Marks Practical Exam 20 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- V)
Title of Paper: Big Data Analytics

Sr. No.	Heading	Particulars
1	**Description the course :**	• The subjects the fundamentals and advanced topics of Big Data Analytics along with use cases.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	4 credits
5	**Hours Allotted :**	60 Hours
6	**Marks Allotted**:	100 Marks
7	**Course Objectives**: • Demonstrates the basic operation in bigdata. • Awareness of Architectures use in bigdata. • Study about Big Data processing language Hadoop.	
8	**Course Outcomes**: • Describes about the modern databases. • Case studies in big data are analyzed.	
9	**Unit I: (20 Hours)** Fundamentals of Big-data analytics, Overview & analytics life cycle, Need, Structured and multi-structured data analysis, Big-data analytics major components, Analytical models and approaches, Relational and non- relational Databases, Application areas, Design and analysis of Analytics model-Analytics design steps, Understanding different data processing models, Statistical models, Predictive models, Descriptive models, Introduction to Data analysis using R- Basics of R Language, Statistical models in R, Statistical techniques applied using R, Graphical analysis techniques, Practical examples/case studies.	
	Unit II (25 Hours) : Analytics Using MapReduce, MapReduce design patterns, Clustered Hadoop environment, Advanced HDFS, Graph Algorithms, Searching and Indexing approaches, MapReduce Applications, Introduction to Pig and HIVE- Programming Pig: Engine for executing data flows in parallel on Hadoop, Programming with Hive: Data warehouse system for Hadoop	
	Unit III (15 Hours) : Case Studies -Real time stream processing using MapReduce and R and Text Analytics based on different documents for identifying interesting patterns and correlations	
10	**Text Books: NA**	
11	**Reference Books:** 8. M. Kantardzic, "Data mining: Concepts, models, methods and algorithms, John Wiley &Sons Inc. 9. Adriaans, Dolf Zantinge , "Data Mining", Pearson Education Asia 10. J. Han, M. Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**

13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 10 Marks Assignment 10 Marks Oral Test 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Syllabus
BSc. (Data Science and Sports Studies)
(Sem.- VI)
Title of Paper: Project 2

Sr. No.	Heading	Particulars
1	**Description the course :**	• The Project Implementation course in Analytics builds upon the knowledge and skills acquired in previous semesters, focusing on the practical implementation of data science projects. • Students will work on real-world projects, applying data science methodologies and techniques to solve complex problems. • They will gain hands-on experience in data collection, preprocessing, feature engineering, and model development.
2	**Vertical :**	Major
3	**Type :**	Theory
4	**Credit**:	2 credits
5	**Hours Allotted :**	30 Hours
6	**Marks Allotted**:	50 Marks
7	**Course Objectives**: • To enable students to apply the knowledge and skills acquired during the B.Sc Data Science program in the implementation of a data science project. • To provide students with hands-on experience in implementing data science methodologies and techniques to solve real-world problems.	
8	**Course Outcomes**: • To develop students' project management skills by planning, executing, and monitoring a data science project. • To enhance students' collaboration and teamwork abilities by working in groups to implement a data science project. • To foster students' critical thinking and problem-solving skills by addressing challenges and making informed decisions during the project implementation phase.	

9	**Unit I: (30 Hours)** Students can undertake industrial project.	
10	**Text Books: NA**	
11	**Reference Books:**	
12	**Internal Continuous Assessment: 40%**	**Semester End Examination: 60%**
13	**Continuous Evaluation through:** Class Tests 10 Marks Presentation 10 Marks Assignment 10 Marks Oral Test 10 Marks	
14	**Format of Question Paper:** for the final examination (Semester End Examination) Question 1 Compulsory 20 Marks Question 2-7 each of 10 Marks (Attempt any 4)	

Letter Grades and Grade Points:

Semester GPA/ Programme CGPA Semester/ Programme	% of Marks	Alpha-Sign/ Letter Grade Result	Grading Point
9.00 - 10.00	90.0 - 100	O (Outstanding)	10
8.00 - < 9.00	80.0 - < 90.0	A+ (Excellent)	9
7.00 - < 8.00	70.0 - < 80.0	A (Very Good)	8
6.00 - < 7.00	60.0 - < 70.0	B+ (Good)	7
5.50 - < 6.00	55.0 - < 60.0	B (Above Average)	6
5.00 - < 5.50	50.0 - < 55.0	C (Average)	5
4.00 - < 5.00	40.0 - < 50.0	P (Pass)	4
Below 4.00	Below 40.0	F (Fail)	0
Ab (Absent)	-	Ab (Absent)	0

Sd/-
Sign of the BOS Chairman
Dr. Manoj N. Reddy
Ad-hoc Board of Studies in Sports Science and Management

Sd/-
Sign of the Offg. Associate Dean
Dr. C.A.Chakradeo
Faculty of Interdisciplinary Studies

Sd/-
Sign of the Offg. Associate Dean
Dr. Kunal Ingle
Faculty of Interdisciplinary Studies

Sd/-
Sign of the Offg. Dean
Prof. A. K. Singh
Faculty of Interdisciplinary Studies