Dr. N.G.P. ARTS AND SCIENCE COLLEGE

(An Autonomous Institution, Affiliated to Bharathiar University, Coimbatore)
Approved by Government of Tamil Nadu and Accredited by NAAC with 'A++' Grade (3rd Cycle-3.64 CGPA)
Dr. N.G.P. - KalapattiRoad, Coimbatore-641048, Tamil Nadu, India
Web: www.drngpasc.ac.in |Email: info@drngpasc.ac.in | Phone: +91-422-2369100

REGULATIONS 2023-24 for Post Graduate Programme (Outcome Based Education model with Choice Based Credit System)

Master of Science in Computer Science with Data Analytics Degree

(For the students admitted during the academic year 2023-24 and onwards)

Programme: M. Sc. (Computer Science with Data Analytics) Eligibility

Candidates for admission to the first year of the **Master of Science (Computer Science with Data Analytics)** Degree Programme shall be required to have passed in B.Sc. Computer Science/ B.C.A. / B.Sc. Computer Technology / B.Sc. Information Technology / B.Sc. Information Sciences / B.Sc. Information Systems / B.Sc. Software Systems / B.Sc. Software Sciences / B.Sc. Applied Sciences (Computer Science / Computer Technology) / B.Sc. Electronics / B.Sc. Mathematics of any University in Tamil Nadu or an Examination accepted as equivalent thereto by the Academic council, subject to conditions as may be prescribed are permitted to appear and qualify for the **Master of Computer Science with Data Analytics Degree Examination** of this College after a programme of study of two academic years.

Programme Educational Objectives

The Curriculum is designed to attain the following learning goals which students shall accomplish by the time of their graduation:

1. Exhibit technical proficiency in Data Analytics to solve real world problems.
2. Engage in successful careers in industry, research and public service.
3. Employ cutting edge tools and technologies for decision making and remain self-motivated and lifelong learners.
4. Practice profession with ethics, integrity, leadership and social responsibility
5. Apply knowledge in areas of Data Analytics for research and entrepreneurship

PROGRAMME OUTCOMES

On the successful completion of the program, the following are the expected outcomes.

PO Number	PO Statement
PO1	Ability to apply knowledge of Computer Science, Mathematics and Statistics to solve problem
PO2	Ability to model, analyze, design, visualize and realize physical systems or processes of increasing size and complexity
PO3	Ability to select appropriate methods and tools for data analysis in specific organizational contexts
PO4	Ability to analyze very large data sets in the context of real world problems and interpret results
PO5	Ability to exhibit soft skills and understand professional and social responsibilities

M.Sc. Computer Science with Data Analytics Credit Distribution

Part	Subjects	No. of Papers	Credit		Semester No.
III	Core	11	3 x 5 = 15 8 x 4 = 32	47	I - III
	Core Practical	06	6 x 2 =12	12	I - III
	DSE	03	03 x 04 = 12		I - III
	EDC	01	01 x 04 = 04		II
	Industrial Training	01	01 x 02 = 02		III
	Core Project	01	01 x 15 = 15		IV
TOTAL CREDITS			**92**		

CURRICULUM

M.Sc. COMPUTER SCIENCE WITH DATA ANALYTICS

Course Code	Course Category	Course Name	L	T	P	Exam (h)	Max Marks			Credits
							CIA	ESE	Total	
First Semester										
234DA2A1CA	Core I	Principles of Data Science and Python	4	1	-	3	25	75	100	5
234DA2A1CB	Core II	Probability and Statistics	4	1	-	3	25	75	100	4
234DA2A1CC	Core III	Design and Analysis of Algorithms	4	-	-	3	25	75	100	4
234CS2A1CB	Core IV	Advanced Java	4	-	-	3	25	75	100	4
234DA2A1CP	Core Practical I	Python Programming	-	-	4	3	40	60	100	2
234CS2A1CQ	Core Practical II	Advanced Java	-	-	4	3	40	60	100	2
234DA2A1DA	DSE I	Digital Image Processing	4	-	-	3	25	75	100	4
234DA2A1DB		Information Retrieval								
234DA2A1DC		Web Intelligence								
Total			20	2	8	-	-	-	700	25

9/6/23
BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science College
APPROVED
8th 9/6/23 | AC - 15th 14/7/23 | GB - 20th 5/8/23

Course Code	Course Category	Course Name	L	T	P	Exam (h)	Max Marks			Credits
							CIA	ESE	Total	
Second Semester										
234DA2A2CA	Core V	Artificial Intelligence	4	1	-	3	25	75	100	5
234DA2A2CB	Core VI	Data Mining	4	1	-	3	25	75	100	4
234DA2A2CC	Core VII	Information and Network Security	4	-	-	3	25	75	100	4
235BA2A2EA	EDC	Digital Marketing Analytics	4	-	-	3	25	75	100	4
234DA2A2CP	Core Practical III	R for Data Analytics	-	-	4	3	40	60	100	2
234DA2A2CQ	Core Practical IV	Data Mining	-	-	4	3	40	60	100	2
234DA2A2DA	DSE II	Customer Analytics	4	-	-	3	25	75	100	4
234DA2A2DB		Natural Language Processing								
234DA2A2DC		Advanced Statistics								
Total			**20**	**2**	**8**	-	-	-	**700**	**25**

11/10/23
BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science College
APPROVED
BoS- 9th 16.10.2023 | AC - 16th 13.12.2023 | GB - 21st 05.01.2024

Academic Council
CHAIRMAN
Dr.N.G.P. Arts and Science College

Course Code	Course Category	Course Name	L	T	P	Exam (h)	Max Marks			Credits
							CIA	ESE	Total	
Third Semester										
234DA2A3CA	Core VIII	Machine Learning	4	1	-	3	25	75	100	5
234DA2A3CB	Core IX	Internet of Things and Applications	4	-	-	3	25	75	100	4
234DA2A3CC	Core X	Cloud Computing	4	-	-	3	25	75	100	4
234DA2A3CD	Core XI	Big Data Analytics	4	1	-	3	25	75	100	4
234DA2A3CP	Core Practical V	Machine Learning	-	-	4	3	40	60	100	2
234DA2A3CQ	Core Practical VI	Big Data Analytics and Visualization	-	-	4	3	40	60	100	2
234DA2A3DA	DSE III	Data Visualization	4	-	-	3	25	75	100	4
234DA2A3DB		Modern Databases								
234DA2A3DC		Deep Learning								
234DA2A3TA	Core XII	Industrial Training	-	-	-	3	50	50	100	2
		Total	**20**	**2**	**8**	-	-	-	**800**	**27**

BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science College
APPROVED
BoS- 10th 2.4.24 | AC - 17th 17.4.24 | GB -

Academic Council CHAIRMAN Dr.N.G.P. Arts and Science College

Course Code	Course Category	Course Name	L	T	P	Exam (h)	Max Marks			Credits
							CIA	ESE	Total	
Fourth	Semester									
234DA2A4CV	Core:XIII	Project and Viva-Voce	-	-	-	3	80	120	200	15
		Total	-	-	-	-	-	-	200	15
							***Grand Total**		**2400**	**92**

***Total Credits does not exceed 92 credits**

BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science Colle...
APPROVED
BoS. 11th 05.11.24 | AC - 18th 26.11.24 | GB -

D.N.G.P. Arts and Science College * Academic Council * CHAIRMAN

DISCIPLINE SPECIFIC ELECTIVE

Students shall select the desired course of their choice in the listed elective course during Semesters I, II & III

Semester I (Elective I)

List of Elective Courses

S. No.	Course Code	Name of the Course
1	234DA2A1DA	Digital Image Processing
2	234DA2A1DB	Information Retrieval
3	234DA2A1DC	Web Intelligence

Semester II (Elective II)

List of Elective Courses

S. No.	Course Code	Name of the Course
1	234DA2A2DA	Customer Analytics
2	234DA2A2DB	Natural Language Processing
3	234DA2A2DC	Advanced Statistics

Semester III (Elective III)

List of Elective Courses

S. No.	Course Code	Name of the Course
1	234DA2A3DA	Data Visualization
2	234DA2A3DB	Modern Databases
3	234DA2A3DC	Deep Learning

EXTRA CREDIT COURSES

The following are the courses offered under self study to earn extra credits

Semester III

S. No.	Course Code	Name of the Course
1	234DA2ASSA	Business Analytics
2	234DA2ASSB	Professional Ethics

PG REGULATION (R5)

(2023-24 and onwards)

(OUTCOME BASED EDUCATION WITH CBCS)

Effective from the academic year 2023-24 and applicable to the students admitted to the Degree of Master of Arts/Commerce/Management/Science.

1.NOMENCLATURE

1.1 Faculty: Refers to a group of programmes concerned with a major division of knowledge. Eg. Faculty of Computer Science consists of Programmes like Computer Science, Information Technology, Computer Technology, Computer Applications, Cognitive Systems, Artificial Intelligence and Machine Learning and Cyber Security and Data Analytics etc.

1.2 Programme: Refers to the Master of Arts/Management/Commerce/Science Stream that a student has chosen for study.

1.3 Batch: Refers to the starting and completion year of a programme of study. Eg. Batch of 2023–2025 refers to students belonging to a 2-year Degree programme admitted in 2023 and completing in 2025.

1.4 Course: Refers to component of a programme. A course may be designed to involve lectures / tutorials / laboratory work / seminar / project work/ practical training / report writing / Viva voce, etc or a combination of these, to effectively meet the teaching and learning needs and the credits may be assigned suitably.

a) Core Courses A course, which should compulsorily be studied by a candidate as a core requirement is termed as a Core course.

b) Extra Departmental Course (EDC): A course chosen generally from a related discipline/subject, with an intention to seek exposure in the discipline relating to the core domain of the student.

c) Discipline Specific Elective Course (DSE): Elective courses are offered under main discipline/ subject of study.

d)Internship/Industrial Training (IT)

Students must undertake industrial / institutional training for a minimum of 15 days during the II semester summer vacation. The students will submit the report for evaluation during III semester.

e) Project Work: It is considered as a special course involving application of knowledge in problem solving/analyzing/exploring a real-life situation. The Project work will be given in lieu of a Core paper.

f) Extra credits Extra credits will be awarded to a student for achievements in co-curricular activities carried out outside the regular class hours. The guidelines for the award of extra credits are given in section two, these credits are not mandatory for completing the programme.

g) Advanced Learner Course (ALC): ALC is doing work of a higher standard than usual for students at that stage in their education. Research work / internships carried out in University/ Research Institutions/ Industries of repute in India or abroad for a period of 15 to 30 days.

2. STRUCTURE OF PROGRAMME

- Core Course
- Extra Departmental Course (EDC)
- Discipline Specific Elective (DSE)
- Industrial Training (IT)
- Project

3. DURATION OF THE PROGRAMME

M.Sc. /M.Com. / M.A. Programme must be completed within 2 Years (4 semesters) and maximum of 4 Years (8 semesters) from the date of acceptance to the programme. If not, the candidate must enroll in the course determined to be an equivalent by BoS in the most recent curriculum recommended for the Programme.

4. REQUIREMENTS FOR COMPLETION OF A SEMESTER

Every student shall ordinarily be allowed to keep terms for the given semester in a program of his/ her enrolment, only if he/ she fulfills at least seventy five percent (75%) of the attendance taken as an average of the total number of lectures,

practicals, tutorials, etc. wherein short and/or long excursions/field visits/study tours organised by the college and supervised by the faculty as envisaged in the syllabus shall be credited to his attendance. Every student shall have a minimum of 75% as an overall attendance.

5. EXAMINATIONS

The end semester examinations shall normally be conducted after completing 90 working days for each semester. The maximum marks for each theory and practical course as follows,

Mark distribution for Theory Courses

Continuous Internal Assessment (CIA) : 40 Marks

End Semester Exams (ESE) : 60 Marks

Total : 100 Marks

i) Distribution of Internal Marks

S.No.	Particulars	Distribution of Marks
1	CIA I (2.5 Units) (On completion of 45th working day)	5
2	Model (All 5 Units) (On completion of 85th working day)	5
3	Attendance	05
4	Library Usage	05
5	Skill Enhancement *	05
	Total	**25**

Breakup for Attendance Marks:

S.No	Attendance Range	Marks Awarded
1	95% and Above	5
2	90% - 94%	4
3	85% - 89%	3
4	80% - 84%	2
5	75% - 79%	1

Note:

Special Cases such as NCC, NSS, Sports, Advanced Learner Course, Summer Fellowship and Medical Conditions etc. the attendance exemption may be given by principal and Mark may be awarded.

Break up for Library Marks:

S.No	Attendance Range	Marks Awarded
1	10h and above	5
2	9h- less than 10h	4
3	8h - less than 9h	3
4	7h - less than 8h	2
5	6h - less than 7h	1

Note:
In exception, the utilization of e-resources of library will be considered.

*Components for "Skill Enhancement" may include the following:

Class Participation, Case Studies Presentation/Term paper, Field Study, Field Survey, Group Discussion, Term Paper, Presentation of Papers in Conferences, Industry Visit, Book Review, Journal Review, e-content Creation, Model Preparation, Seminar and Assignment.

Components for Skill Enhancement
Any one of the following should be selected by the course coordinator

S.No.	Skill Enhancement	Description
1	Class Participation	• Engagement in class • Listening Skills • Behaviour
2	Case Study Presentation/ Term Paper	• Identification of the problem • Case Analysis • Effective Solution using creativity/imagination
3	Field Study	• Selection of Topic • Demonstration of Topic • Analysis & Conclusion
4	Field Survey	• Chosen Problem • Design and quality of survey • Analysis of survey
5	Group Discussion	• Communication skills • Subject knowledge • Attitude and way of presentation • Confidence • Listening Skill
6	Presentation of Papers in Conferences	• Sponsored • International/National • Presentation • Report Submission

7	Industry Visit	• Chosen Domain • Quality of the work • Analysis of the Report • Presentation
8	Book Review	• Content • Interpretation and Inferences of the text • Supporting Details • Presentation
9	Journal Review	• Analytical Thinking • Interpretation and Inferences • Exploring the perception if chosen genre • Presentation
10	e-content Creation	• Logo/ Tagline • Purpose • Content (Writing, designing and posting in Social Media) • Presentation
11	Model Preparation	• Theme/ Topic • Depth of background Knowledge • Creativity • Presentation
12	Seminar	• Knowledge and Content • Organization • Understanding • Presentation
13	Assignment	• Content and Style • Spelling and Grammar • References

ii) Distribution of External Marks

Total	:	75
Written Exam	:	75

Marks Distribution for Practical course

Total	:	100
Internal	:	40
External	:	60

i) Distribution of Internals Marks

S. No.	Particulars	Distribution of Marks
1	Experiments/Exercises	15
2	Test 1	10
3	Test 2	10
4	Observation Notebook	05
	Total	40

ii) Distribution of Externals Marks

S.No.	Particulars	External Marks
1	Practical	40
2	Record	10
3	Viva- voce	10
	Total	60

Practical examination shall be evaluated jointly by Internal and External Examiners.

A) Mark Distribution for Project

Total	:	200
Internal	:	80
External	:	120

i) Distribution of Internal Marks

S.No.	Particulars	Internal Marks
1	Review I	30
2	Review II	40
3	Attendance	10
	Total	80

ii) Distribution of External Marks

S.No	Particulars	External Marks
1	Project Work & Presentation	100
2	Viva -voce	20
	Total	120

Evaluation of Project Work shall be done jointly by Internal and External Examiners.

6 . Credit Transfer

a. Upon successful completion of 1 NPTEL Course (4 Credit Course) recommended by the department, during Semester I to II, a student shall be eligible to get exemption of one 4 **credit course** during the 3rd semester. The proposed NPTEL course should cover content/syllabus of exempted core paper in 3rd semester.

S. No.	Course Code	Course Name	Proposed NPTEL Course	Credit
1			Option – 1 Paper title	4
			Option – 2 Paper title	
			Option – 3 Paper title	

b. Upon successful completion of **2 NPTEL Courses** (2 Credit each) recommended by the department, during Semester I to II, a student shall be eligible to get exemption of **one 4 credit course** during the 3rd semester. Out of 2 NPTEL proposed courses, **at least 1 course** should cover content/syllabus of exempted core paper in 3rd semester.

Mandatory

The exempted core paper in the 3rd semester should be submitted by the students for approval before the end of 2nd semester

Credit transfer will be decided by equivalence committee

S. No.	Course Code	Course Name	Proposed NPTEL Course	Credit
1			Option – 1 Paper title	2
			Option – 2 Paper title	
			Option – 3 Paper title	
2			Option – 1 Paper title	2
			Option – 2 Paper title	
			Option – 3 Paper title	

NPTEL Courses to be carried out during semester I – II.

S. No.	Student Name	Class	Proposed NPTEL Course		Proposed Course for Exemption
			Course I	Option 1- Paper Title Option 2- Paper Title Option 3- Paper Title	Any one Core Paper in 3rd Semester
			Course II	Option 1- Paper Title Option 2- Paper Title Option 3- Paper Title	

Class Advisor HoD Dean

Mark Distribution for Internship/ Industrial Training

Total : 100
Internal : 40
External : 60

i) Distribution of Internal Marks

S.No.	Particulars	Internal Marks
1	Review I	15
2	Review II	20
3	Attendance	5
	Total	40

ii) Distribution of External Marks

S.No	Particulars	External Marks
1	Internship /Industrial training Presentation	40
2	Viva –voce	20
	Total	60

Internship/ Industrial training shall be evaluated jointly by Internal and External Examiners.

9. Extra Credits: 10

Earning extra credit is not essential for programme completion. Student is entitled to earn extra credit for achievement in Curricular/Co-Curricular/ Extracurricular activities carried out other than the regular class hours.

A student is permitted to earn a maximum of 10 extra Credits during the programme period.
A maximum of 1 credit under each category is permissible.

Category	Credit
Self study Course	1
CA/ICSI/CMA (Foundations)	1
CA/ICSI/CMA (Inter)	1
Sports and Games	1
Publications / Conference Presentations (Oral/Poster)/Awards	1
Innovation / Incubation / Patent / Sponsored Projects / Consultancy	1
Representation in State / National level celebrations	1
Awards/Recognitions/Fellowships	1
Advanced Learner Course (ALC)*	2

Credit shall be awarded for achievements of the student during the period of study only.

GUIDELINES

Self study Course

A pass in the self study courses offered by the department.

The candidate should register the self study course offered by the department only in the III semester.

CA/ICSI/CMA(Foundations)

Qualifying foundation in CA/ICSI/CMA / etc.

CA/ICSI/CMA(Inter)

Qualifying Inter in CA/ICSI/CMA / etc.

Sports and Games

The Student can earn extra credit based on their Achievement in sports in University/ State / National/ International.

Publications / Conference Presentations (Oral/Poster)

Research Publications in Journals

Oral/Poster presentation in Conference

Innovation / Incubation / Patent / Sponsored Projects / Consultancy

Development of model/ Products /Prototype /Process/App/Registration of Patents/ Copyrights/Trademarks/Sponsored Projects /Consultancy

Representation in State/ National level celebrations

State / National level celebrations such as Independence day, Republic day Parade, National Integration camp etc.

Awards/Recognitions/Fellowships

Regional/ State / National level awards/ Recognitions/Fellowships

***Advanced Learner Course (ALC):**

ALC is doing work of a higher standard than usual for students at that stage in their education.

Research work/internships carried out in University/ Research Institutions/ Industries of repute in India or abroad for a period of 15 to 30 days will be considered as Advanced Learners Course.

QUESTION PAPER PATTERN

CIA Test I : [1½ Hours-2.5 Units] - 25 Marks

SECTION	MARKS	DESCRIPTION	TOTAL	Remarks
Section – A	8 x 0.5= 04 Marks	MCQ	25 Marks	Marks secured will be converted To 5 mark
Section - B	3 x 2 = 06 Marks	Answer ALL Questions Either or Type ALL Questions Carry Equal Marks		
Section - C	3 x 05 = 15 Marks			

CIA Test II/ Model [3 Hours-5 Units] - 75 Marks

SECTION	MARKS	DESCRIPTION	TOTAL	Remarks
Section – A	10 x 1 = 10 Marks	MCQ		
Section - B	5 x 3 = 15 Marks	Answer ALL Questions (Either or Type Questions) Each Questions Carry Equal Marks	75 Marks	Marks secured will be converted To 5 mark
Section - C	5 x 8 = 40 Marks			
Section - D	1 x 10 = 10 Marks	Compulsory Question		

End Semester Examination [3 Hours-5 Units] - 75 Marks

SECTION	MARKS	DESCRIPTION	TOTAL
Section – A	10 x 1 = 10 Marks	MCQ	
Section - B	5 x 3 = 15 Marks	Answer ALL Questions (Either or Type Questions) Each Questions Carry Equal Marks	75 Marks
Section - C	5 x 8 = 40 Marks		
Section - D	1 x 10 = 10 Marks	Compulsory Question	

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1CA	Principles of Data Science and Python	CORE	4	1	-	5

PREAMBLE

This course has been designed for students to learn and understand

- Concepts of Data Science
- Understand about Python Programming
- Plotting and Visualization in Python

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the principles of data science	K2
CO2	Understand the techniques for Data Handling	K2
CO3	Apply Numpy and Pandas to perform numerical operations	K3
CO4	Apply the concepts of Python for Data Aggregation and Wrangling	K3
CO5	Create the visualization concepts in Python	K6

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1CA	PRINCIPLES OF DATA SCIENCE AND PYTHON	SEMESTER I

Total Credits: 5

Total Instruction Hours: 60 h

Syllabus

Unit I Data science 12 h

Introduction: Benefits of Data Science - Facets of Data –Big data eco system and data science - Data science process: Steps in data science process - Retrieving data – Data preparation – Data exploration – Data modeling – Presentation - Case Study

Unit II Handling Large Data 12 h

Problems when handling large data – General techniques for handling large data – General techniques for handling large volumes of data–General Programming dealing with Large Data Sets – Steps in big data – Distributing data storage and processing with Frameworks- Applications in Data Science - Case Study – Assessing risk when loaning money

Unit III Numpy and Pandas 12 h

Introduction to NumPy - Understanding the N - dimensional data structure - Creating NumPy arrays - Basic operations and manipulations on N-dimensional arrays - Indexing and Slicing-Advanced Indexing – Pandas: Mathematical Functions – Statistical Functions – Search, Sorting and Counting Functions –Matrix Library

Unit IV Data Aggregation and Operations 12 h

Introduction: GroupBy Mechanics – Data Aggregation – Groupwise Operations and Transformations – Pivot Tables and Cross Tabulations – Date and Time Date - Type tools – Time Series Basics – Data Ranges - Frequencies and Shifting Combining and Merging DataSets – Reshaping and Pivoting – Data Transformation – String Manipulation, Regular Expressions

Unit V Plotting and Visualization 12 h

Introduction: Data Acquisition by Scraping web applications –Submitting a form - Fetching web pages – CSS Selectors. Visualization: Visualization In Python: Matplotlib package – Plotting Graphs – Controlling Graph – Adding Text – More Graph Types – Getting and setting values - Plotting with Pandas and seaborn-Line plots-Bar Plots-Histogram Density and Plots - scatter or point plots-facet grids and categorical data

Text Books

1 Davy Cielen, Arno D.B. Meysmen, Mohamed Ali, 2020, "Introducing Data Science", Dream Tech Press (UNITS I,II)

2 Wes Mc Kinney, 2020, "Python for Data Analysis", 5th Edition, O'Reilly (UNITS III, IV, V)

References

1 John V Guttag, 2016, "Introduction to Computation and Programming Using Python", 2nd Edition., MIT press]

2 Gypsy Nandi, Rupam Kumar Sharma, 2020, "Data Science Fundamentals and Practical Approach, BPB

3 Zed Shaw, 2014, "Learn Python the Hard Way", 3rd Edition, Addison-Wesley, USA,

4 Fabio Nelli, 2018, "Python Data Analytics", Second Edition, Apress, New York,

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1CB	PROBABILITY AND STATISTICS	CORE	4	1	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Data numerically and visually
- The knowledge of testing of hypothesis for small and large samples which plays an important role in real life applications
- Data-based claims and quantitative arguments

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Make use of the concepts of probability which can describe real life phenomenon	K2
CO2	Apply discrete and continuous probability distributions in the relevant application areas	K3
CO3	Learn how to develop correlation and regression model and apply for the specific perspective data in appropriate manner	K3
CO4	Analyse a best estimator with reference to the different criteria in case of real-life applications	K4
CO5	Learn the details and complexities of Analysis of Variance (ANOVA)	K4

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓				
CO2	✓		✓	✓	
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development	✓	Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1CB	PROBABILITY AND STATISTICS	SEMESTER I

Total Credits: 4

Total Instruction Hours: 60 h

Syllabus

Unit I Probability Concepts 10 h

Introduction -Probability Defined -Importance of the Concept of Probability - Calculation of Probability -Theorems of Probability -Addition Theorem - Multiplication Theorem -Conditional Probability -Bayes' Theorem -Mathematical Expectation

Unit II Probability Distributions 10 h

Introduction - Binomial Distribution-Fitting a Binomial Distribution- Poisson Distribution - Fitting a Poisson Distribution- Normal Distributions - Fitting a Normal Curve

Unit III Correlation and Regression Analysis 12 h

Correlation- Scatter Diagram Method -Graphic Method- Karl Pearson's Coefficient of Correlation-Spearman's coefficient of Correlation - Regression Analysis - Regression Lines -Regression Equations -Regression Equation of Y on X -Regression Equation of X on Y

Unit IV Statistical Inference -Testing of Hypothesis 14 h

Introduction - Hypothesis Testing - Standard Error and Sampling Distribution - Estimation -Tests of Significance for Large Samples -Difference between small and large samples -Two tailed test for difference between the means of two samples - Standard Error of the difference between two standard deviations -Tests of significance for small samples - Assumption of Normality -Student's t distribution - Application of the t Distribution

Unit V Chi square test, F-test and Analysis of Variance 14 h

Introduction- Chi-Square test- F-Test -Applications of F-Test -Analysis of Variance - Assumptions -Technique of Analysis of Variance - One-Way Classification - Analysis of Variance in Two-Way Classification Model

Text Books

1 Gupta S.P, 2017, "Statistical Methods", 45th Edition, Sultan Chand and Sons, New Delhi

References

1 Ronald E. Walpole,2018," Probability and Statistics", 9th Edition, Pearson Education, South Asi

2 Sheldon M. Ross, 2017, "Introductory Statistics", 4th Edition, Academic Press, United States

3 Vijay K. Rohatgi A.K, MD. Ehsanes Saleh, 2015, " An introduction to Probability and Statistics", 3rd Edition, John Wiley and Sons, New Delhi

4 Sheldon M. Ross, 2017, "A first course in Probability", 5th Edition, PHI, New Jersey

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1CC	DESIGN AND ANALYSIS OF ALGORITHMS	CORE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Design and analysis of algorithm techniques
- Analyze the efficiency of different algorithmic solutions
- Implementation and evaluation of complex algorithms

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the fundamentals of algorithms and data structures	K2
CO2	Apply Divide and Conquer approach using various sorting algorithms	K3
CO3	Analyze Greedy algorithm design technique and its applications	K4
CO4	Interpret Dynamic Programming paradigms to solve real-world problems	K2
CO5	Implement Backtracking, Branch and Bound techniques to solve complex problems	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		
CO2	✓	✓	✓	✓	✓
CO3	✓	✓	✓		
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1CC	DESIGN AND ANALYSIS OF ALGORITHMS	SEMESTER I

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction 8 h

Algorithm Definition - Analyzing and Designing algorithms - Performance Analysis-Asymptotic Notations - Time and Space complexity of an algorithm using O Notation. Elementary Data Structures: Stacks and Queues – Linked lists.

Unit II Divide and Conquer 10 h

Introduction: Strassen's Algorithm for Matrix Multiplication - Sorting and Order Statistics: Heap sort - Algorithm - Priority Queues - Quick Sort - Description, Performance and Analysis – Merge sort.

Unit III The Greedy Method 10 h

The General Method - Knapsack Problem - Minimum Cost Spanning Trees – Prim's Algorithm – Kruskal's Algorithm - Optimal Storage On Tapes – Optimal Merge Patterns - Single Source Shortest Paths – Dijkstra's Algorithm.

Unit IV Dynamic Programming 10 h

The General Method – All-Pairs Shortest Paths – Warshall's and Floyd's Algorithm – Single-Source Shortest Paths - Bellman–Ford Algorithm - Optimal Binary Search Trees - 0/1 Knapsack - Reliability Design - The Traveling Salesperson Problem.

Unit V Backtracking 10 h

The General Method – The 8-Queens Problem – Sumof Subsets– Graph Coloring - Hamiltonian Cycles –Branch and Bound: Knapsack Problem – Travelling Salesman Problem.

Text Books

1. Thomas H. Cormen, Charles E. Leiserson and Ronald L. Rivest, 2009, "Introduction to Algorithms", 3rd Edition, MIT Press
2. Ellis Horowitz, Sartaj Sahni and Sanguthevar Rajasekaran, 2009, "Fundamentals of Computer Algorithms, 2nd Edition, University Press

References

1. Robert L. Kruse and Clovis L. Tondo, 2007, "Data Structures and Program design in C", 2nd Edition, Pearson Education
2. Michael T. Goodrich, Roberto Tamassia",2001,Algorithm Design, Foundations, Analysis, and Internet Examples", 1st Edition., Wiley
3. Mark Allen Weiss,2013, "Introduction to the Design Data Structures and Algorithm Analysis in C++",4th Edition., Addison-Wesley
4. Tim Roughgarden. 2017, "Algorithms Illuminated", Kindle Edition Soundlikeyourself Publishing, New York.

Course Code	Course Name	Category	L	T	P	Credit
234CS2A1CB	ADVANCED JAVA	CORE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Advance Java concepts to develop applications
- The Concepts of Java Beans and Swings
- Database Connectivity using JDBC and Embedded SQL

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand about Java beans and swing	K2
CO2	Understand the life cycle of Java Servlet	K2
CO3	Develop and apply event in JSP and RMI	K3
CO4	Learn the architecture and design of Enterprise Java Bean	K2
CO5	Design applications implementing Database Connectivity using JDBC and Embedded SQL.	K6

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		
CO2	✓	✓	✓		
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓		
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234CS2A1CB	ADVANCED JAVA	SEMESTER I

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Java Beans and Swings 10 h

Introduction: Advantages – Design patterns for Properties – Events – Methods and Design Patterns - Java Beans API – Swing : Introduction – Swing Is Built on the AWT - Two Key features of Swing – MVC Connections – Components and Containers – The Swing Packages – Simple Swing Applications - Exploring Swing

Unit II Java Servlet 10 h

Introduction: Background - The life cycle of a Servlet – Using Tomcat for Servlet development – A Simple Servlet – The Javax.Servlet Packages – Reading Servlet Parameters – The javax.servlet.http packages – Handling Http request and responses – cookies - Session Tracking

Unit III Java Server Pages, Remote Method Invocation 8 h

Java Server Pages- Introduction - Tags: Variable Objects - Request String: Parsing Other Information - User Session - Cookies - Session objects. Java Remote method Invocation: Remote Interface - Passing Objects - RMI Process - Server side - Client side

Unit IV Enterprise Java Bean 10 h

Enterprise Java Beans : The EJB Container – EJB Classes - EJB Interfaces – Deployment Descriptors : Referencing EJB - Sharing Resources - Security Elements - Query Elements - Assembly Elements - Session Java Bean: Stateless and Stateful - Creating a Session Java Bean - Entity Java Bean - Message -Driven Bean

Unit V Database Connectivity 10 h

JDBC Objects : The Concept of JDBC - JDBC Driver types – JDBC Packages – Database Connection – Statement Objects – ResultSet – Transaction Processing - JDBC and Embedded SQL : Tables and Indexing - Inserting, Selecting and Updating Data

Text Books

1. Herbert Schildt, 2018, "Java The Complete Reference", 10th Edition, Tata McGraw Hill (Unit I-II)
2. Jim Keogh, 2002, "J2EE: The Complete Reference", McGraw Hill Education (Unit III - V)

References

1. Herbert Schildt, 2018, "Java, A Beginner Guide", 8th Edn., Oracle Press
2. Bert Bates, KarthySierra, Eric Freeman, Elisabeth Robson, 2009, "Head First Design Patterns", 1st Edition.), O'Reilly
3. Robert Pattinson, 2018, "The Ultimate Beginners Guide for Advance Java" First Edition, Amazon Digital Services LLC
4. E RamarajP Geetha S Muthukumaran, 2018, "Advanced JAVA Programming", 1st Edition, Pearson.,Noida.

234DA2A1CP	PYTHON PROGRAMMING	SEMESTER I

Total Credits: 2
Total Instructions Hours: 48 h

S.No	List of Programs
1	Programs to perform aggregation operations
2	Programs to Implement a sequential search
3	Programs to Explore string functions
4	Programs to Read and Write into a file
5	Programs to Demonstrate use of List
6	Programs to Demonstrate use of Dictionaries
7	Programs to Demonstrate use of Tuples
8	Programs to Create Comma Separate Files (CSV), Load CSV files into internal Data
9	Programs using Pandas: Extract items at given positions from a series
10	Programs to implement correlation and covariance
11	Program to plot graphs using Matplotlib and seaborn packages
12	Programs to Perform Analysis for given data set using Pandas

Note: Ten Programs are mandatory

234CS2A1CQ	ADVANCED JAVA	SEMESTER I

Total Credits: 2
Total Instructions Hours: 48h

S.No	List of Programs
1	Programs using Java control statements.
2	Programs to implement the Collection with Iterator.
3	Programs to create applet incorporating features such as images, shapes, background, and foreground color
4	Create applications using simple GUI
5	Programs to perform some applications using Java Bean
6	Create applications using Swing
7	Programs to demonstrate AWT Components with Event Handling.
8	Programs to perform Session Tracking.
9	Java servlet programs to implement sendredirect () Method (using Http servlet class).
10	Servlet programs using HTTP Servlet.
11	Create web applications using JSP.
12	Programs with JDBC to interact with database.

Note: Ten Programs are mandatory.

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1DA	DIGITAL IMAGE PROCESSING	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Digital image processing fundamentals, color models and image filtering
- Image edge detection and image compression concepts and implement them
- Image segmentation and morphological concepts and implement them

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand image processing fundamentals, its models and color models	K2
CO2	Discuss images filtering concepts and techniques	K2
CO3	Demonstrate image edge detection techniques and applications	K3
CO4	Apply image compression methods and models for real life problems	K3
CO5	Analyze segmentation and morphological image processing	K4

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		
CO2	✓	✓	✓		
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1DA	DIGITAL IMAGE PROCESSING	SEMESTER I

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Fundamentals and Color Models 8 h

Fundamentals: Image Sensing and Acquisition - Image Sampling and Quantization - relationship between Pixels - Random noise - Gaussian Markov Random Field - σ-field, Linear and Non-linear Operations - Image processing models: Causal - Semi-causal - non-causal models - Color Models: Color Fundamentals - Color Models - Pseudo-color Image Processing - Full Color Image Processing - Color Transformation - Noise in Color Images.

Unit II Spatial Domain and Frequency Domain 10 h

Spatial Domain: Enhancement in spatial domain: Point processing - Mask processing - Smoothing Spatial Filters - Sharpening Spatial Filters - Combining Spatial Enhancement Methods - Frequency Domain: Image transforms: FFT - DCT - Karhunen-Loeve transform - Hotlling's T square transform - Wavelet transforms and their properties - Image filtering in frequency domain.

Unit III Edge Detection 10 h

Edge Detection: Types of edges - threshold - zero-crossing - Gradient operators: Roberts - Prewitt - and Sobel operators - residual analysis-based technique - Canny edge detection - Edge features and their applications.

Unit IV Image Compression 10 h

Image Compression: Fundamentals - Image Compression Models - Elements of Information Theory - Error Free Compression: Huff-man coding - Arithmetic coding; - Wavelet transform based coding - Lossy Compression: FFT - DCT - KLT - DPCM - MRFM based compression - Wavelet transform based - Image Compression standards.

Unit V Image Segmentation 10 h

Image Segmentation: Detection and Discontinuities: Edge Linking and Boundary Deduction - Threshold - Region-Based Segmentation - Segmentation by Morphological watersheds - The use of motion in segmentation - Image Segmentation based on Color - Case study.

Morphological Image Processing: Erosion and Dilation, Opening and Closing, Hit-Or-Miss Transformation, Basic Morphological Algorithms, Gray-Scale Morphology

Text Books

1 Rafael Gonzalez, Richard E. Woods, 2019, "Digital Image Processing", (Fourth Edition), Pearson Education (UNIT I, II,IV,V)

2 A. K. Jain, 2015, "Fundamentals of Image Processing", Second Edition, Pearson Education (UNIT III)

References

1 S Annadurai, R Shanmugalakshmi, 2007, " Fundamentals of Digital Image Processing, (First Edition), Pearson Education

2 Todd R.Reed,2015, "Digital Image Sequence Processing, Compression and Analysis", (Sixth Edition), ECRC Press

3 Prasad, S.S.Iyengar, 2015 "Wavelet Analysis with Applications to Image Processing",(Seventh Edition) CRC Press

4 William K. Pratt, 2002, "Digital Image Processing", John Wiley, New York,.

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1DB	INFORMATION RETRIEVAL	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- The concepts of information retrieval techniques
- The techniques focused on document classification, tolerant retrieval and evaluation
- The methods of developing an information retrieval system

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the concepts of the standard models of Information Retrieval	K2
CO2	Understand the methods for handling wild card queries and spelling correction	K2
CO3	Apply appropriate methods for scoring and evaluating IR systems	K3
CO4	Apply text classification to locate relevant information from large collections of text data	K3
CO5	Design an Information Retrieval System for search tasks involving XML and web data	K6

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1DB	INFORMATION RETRIEVAL	SEMESTER I

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction to Information Retrieval 9 h

Introduction to Information Retrieval- Building an inverted Index - Processing Boolean Queries - Boolean Model vs Ranked Retrieval - Term Vocabulary and Postings: Tokenization - Stop words - Normalization-Stemming - Skip pointers - Phrase queries: Biword indexes- Positional indexes

Unit II Tolerant Retrieval, Index construction 10 h

Search Structures for Dictionaries- Wild card queries – General wild card queries- kgram indexes for wild card queries - Spelling correction – Forms- Edit distance - kgram indexes for spelling correction - Phonetic Correction – Index construction- Distributed indexing- Statistical properties of terms : Heaps' Law- Zipf's Law

Unit III Evaluation in Information Retrieval 10 h

Term frequency and weighting –Inverse document frequency- TF-IDF weighting - Vector space model for scoring -Efficient scoring and ranking -Evaluation: Information retrieval system evaluation- Evaluation of unranked retrieval sets- Evaluation of ranked retrieval sets- Case study

Unit IV Text Classification 10 h

Text classification and Naive Bayes- The text classification problem- Naive Bayes text classification – Feature selection – Mutual information- Vector space classification: Document representations and measures of relatedness in vector spaces- k nearest neighbour - Linear versus nonlinear classifiers - Case study

Unit V XML Retrieval and Web search 9 h

XML Indexing and Search: Basic XML concepts - Challenges in XML retrieval- A vector space model for XML retrieval - Data vs. Text-centric XML- Web search basics- Web characteristics-Web crawling – Features of web crawler-Architecture- Distributing indexes indexes - Machine learning methods in ad hoc information retrieval - Case study

Text Books

1 Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schuetze, 2009 , "Introduction to Information Retrieval", Edition, Cambridge University Press

References

1 Baeza -Yates Ricardo and Berthier Ribeiro - Neto, 2011, "Modern Information Retrieval",. 2nd edition, Addison-Wesley

2 Gerald Kowalski, 2010, "Information Retrieval Architecture and Algorithms", First Edition, Berlin, Heidelberg: Springer-Verlag

3 G.G. Chowdhury, 2010, "Introduction to Modern Information Retrieval", 3rd Edition, Facet Publishing

4 Bruce Croft, Donald Metzler, and Trevor Strohman, , 2009, "Search Engines: Information Retrieval in Practice" Pearson Education

Course Code	Course Name	Category	L	T	P	Credit
234DA2A1DC	WEB INTELLIGENCE	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- The concepts of web mining and crawling
- The techniques in opinion mining and sentiment analysis
- The concepts of social network Analysis

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the concepts of web mining	K2
CO2	Analyze social networks and web crawling	K4
CO3	Experiment with opinion mining and sentiment analysis	K5
CO4	Understand Google Analytics	K2
CO5	Design Applications using web intelligence	K5

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A1DC	WEB INTELLIGENCE	SEMESTER I

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Fundamentals of Web Mining 10 h

Introduction - Web Mining: Information Retrieval and Web Search - Basic Concepts of Information Retrieval - Information Retrieval Models - Relevance Feedback - Evaluation Measures - Text and Web Page Pre-Processing - Web Search - Meta Search: Combining Multiple Rankings - Web Spamming

Unit II Social Network Analysis and Web Crawling 10 h

Social Network Analysis - Co-Citation and Bibliographic Coupling - Page Rank - Semantic web - Web Intelligence: Levels - Goals - Characteristics - Challenges and issues Tools for web crawling - Web Crawling: Basic Crawler Algorithm - Implementation Issues - Universal Crawlers - Focused Crawlers - Topical Crawlers: Topical Locality and Cues - Best-First Variations - Adaptation - Evaluation - Crawler Ethics and Conflicts

Unit III Opinion Mining and Sentiment Analysis 8 h

The Problem of Opinion Mining - Document Sentiment Classification - Sentence Subjectivity and Sentiment Classification – Opinion Lexicon Expansion - Aspect-Based Opinion Mining – Mining Comparative Opinions - Opinion Search and Retrieval - Case study

Unit IV Google Analytics 10 h

Google Analytics: Introduction - Cookies - Accounts vs Property - Tracking Code - Tracking Unique Visitors - Demographics - Page Views and Bounce Rate Acquisitions -Custom Reporting - Case study

Unit V Applications 10 h

Applications: Filters - Ecommerce Tracking - Real Time Reports - Customer Data-Alert - Adwords Linking - Adsense Linking - Attribution Modeling - Segmentation - Campaign Tracking - Multi-Channel Attribution - Case Study - Recommendation engines based on users, items and contents

Text Books

1. Bing Liu ,2011, "Web Data Mining Exploring Hyperlinks, Contents, and Usage Data", 2nd Edition, Springer(Unit I-III)
2. Ning Zhong , Jiming Liu and Yiyu Yao, 2010, "Web Intelligence", Springer(Unit IV,V)

References

1. Ricardo Baeza -Yates and BerthierRibeiro-Neto,2011,"Information Retrieval: The Concepts and Technology behind Search,2nd Edition, ACM Press
2. Juan D. Velasquez, Lakhmi C. Jain (Eds.),2010,"Advanced Techniques in Web Intelligence - 1",1st Edition, Springer
3. Mark Levene,2010,"An Introduction to Search Engines and Web Navigation",2nd Edition, Wiley
4. Eric Fettman, Shiraz Asif, Feras Alhlou , 2016 "Google Analytics Breakthrough", Wiley

8/6/23

BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science Col...
APPROVED

BoS - 8th	AC - 15th	GB - 20th
9/6/23	14/7/23	5/8/23

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2CA	ARTIFICIAL INTELLIGENCE	CORE	4	1	-	5

PREAMBLE

This course has been designed for students to learn and understand

- The principles and concepts of Artificial Intelligence
- The various AI approaches towards problem solving, knowledge representation, and reasoning
- The applications of AI and analytics in IoT, cryptocurrency, chatbots and connected cars

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the concepts of Artificial Intelligence	K2
CO2	Analyze a suitable search method for real world problem	K4
CO3	Apply Constraint Satisfaction techniques and Adversarial search methods	K3
CO4	Understand knowledge representation and reasoning in AI	K2
CO5	Discuss the applications of AI and Analytics in various domains	K2

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	✓
CO3	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2CA	ARTIFICIAL INTELLIGENCE	SEMESTER II

Total Credits: 5

Total Instruction Hours: 60 h

Syllabus

Unit I Intelligent Agents and Search Strategies 12 h

Foundations of Artificial Intelligence - Intelligent Agents: Agents and Environments- Structure of Agents – Problem Solving: Problem Solving Agents- Problem Formulation - Uninformed Search Strategies: Breadth-First Search – Depth-First Search – Depth-Limited Search – Iterative Deepening Depth-First Search – Bidirectional Search - Comparing Uninformed Search Strategies- Case Study on Intelligent Agents

Unit II Heuristic and Classical Search 12 h

Greedy Best-First Search – A * Search – Memory-Bounded Heuristic Search – Heuristic Functions – Local Search Algorithms – Hill Climbing Search – Simulated Annealing –Local Beam Search – Genetic Algorithms – Online Search Agents and Unknown Environments: Online Search Problems - Online Search Agents - Online Local Search –Learning in Online Search- Case Study on Online Search Agents

Unit III Constraint Satisfaction Problems 12 h

Introduction : Constraint Satisfaction Problems(CSP): Backtracking search for CSP- Local Search for CSP - Structure of Problems - Adversarial Search: Introduction – Games – Optimal Decision in Games - The Min Max Algorithm – Alpha–Beta Pruning – Games that Include an Element of Chance: Card Games- Case Study on Optimal Decision in Games

Unit IV Knowledge Representation and Reasoning 12 h

Knowledge Based Agents – Logic – Propositional Logic: Syntax – Semantics – A simple knowledge base – Inference – Equivalence, Validity and Satisfiability – Reasoning Patterns in Propositional Logic: Resolution – Forward and Backward Chaining - First Order Logic : Syntax and Semantics of First Order Logic – Using First Order Logic - Case Study on Knowledge Based Agents

Unit V Applications of Analytics and AI 12 h

Exponential Technologies Underpinned by Analytics and AI : Beating Cyberattacks with Analytics-Connected Car Technology Reshaping Automotive Industry-IoT

Analytics: Extracting Value and Transforming Business- Cryptocurrency Analytics: Deep Insights into the new Asset Class -Chatbots: The Protege of AI and Analytics- Case Study on Artificial Intelligence in Tesla

Text Books

1. Russell, S.J. and Norvig, P.,2015, "Artificial Intelligence: A Modern Approach", 3rd Edition, Pearson Education [Units I-III]
2. Sameer Dhanrajani., 2018, : AI and Analytics Accelerating Business Decisions", 1st Edition, Wiley [Units IV-V]

References

1. Lavika Goel,., 2021, " Artificial Intelligence: concepts and Applications", 1st Edition, Wiley
2. Patterson, D.W., 2012,"Introduction to Artificial Intelligence and Expert Systems", Prentice Hall of India
3. Nilsson, N.J, 2011,"Artificial Intelligence A New Synthesis", 1st Edition, Elsevier
4. Kevin Kight, Elaine Rich, 2017, Artificial Intelligence, 3rd Edition, McGrawHill Publishing

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2CB	DATA MINING	CORE	4	1	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Be familiar with foundations of data mining and Data Preprocessing
- Understand algorithms in data mining for data classification and clustering
- Know Data Mining Trends and Research

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the types of the data to be mined and present a general classification of tasks and primitives to integrate a data mining system.	K2
CO2	Understand interesting patterns from large amounts of data using classification.	K2
CO3	Discover the role played by Clustering in data mining.	K4
CO4	Choose and employ suitable data mining algorithms to build association rules.	K5
CO5	Evaluate the scope of data mining tools and Techniques	K5

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	✓
CO3	✓	✓	✓	✓	✓
CO4	✓		✓	✓	✓
CO5	✓	✓	✓	✓	✓

COURSE FOCUSES ON

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2CB	DATA MINING	SEMESTER II

Total Credits: 4

Total Instruction Hours: 60 h

Syllabus

Unit I Introduction to Data Mining 12 h

Introduction: Data Mining - Database Data – Data Mining Technologies - Data Objects and Attribute types - Measuring Data Similarity and Dissimilarity: Proximity Measures for Nominal Attributes - Proximity Measures for Binary Attributes - Minkowski Distance - Data Preprocessing: Data Cleaning - Data Integration - Data Reduction: Wavelet Transforms - Principal Components Analysis - Regression Models - Data Transformation by Normalization - Discretization: Binning - Histogram

Unit II Classification 12 h

Introduction - Decision Tree Induction: Attribute Selection Measures - Tree Pruning - Bayes Classification Methods: Naïve Bayesian Classification - Rule Based Classification - Model Evaluation and Selection: Metrics for Evaluating Classifier performance - Techniques to improve Classification accuracy: Ensamble Methods - Bagging - Boosting - Random Forests - Advanced Methods: Support Vector Machines - K-Nearest Neighbor Classifiers

Unit III Cluster Analysis 12 h

Introduction - Partitioning Methods - Hierarchical Methods: Agglomerative - Hierarchical - Distance measures in Algorithmic methods -BIRCH - Density Based Methods: DBSCAN - Grid Based Methods: STING - Evaluation of Clustering: Determining the Number of Clusters - Measuring Clustering Quality - Clustering high dimensional data: Problems - Challenges - Methodologies - Clustering graph and network data.

Unit IV Outlier Analysis, Mining Frequent Patterns, Associations and Correlations 12 h

Outlier Detection: Outlier and Outlier analysis - statistical approaches - proximity based approaches - Clustering based approaches - Mining Frequent Patterns, Associations and Correlations: Market Basket Analysis - Apriori Algorithm - Association Rules from frequent itemsets - Advanced Pattern Mining: Pattern Mining in Multilevel, Multidimensional Space

Unit V Data Mining Trends and Research 12 h

Mining Complex data types - Other methodologies of data mining - Data Mining Applications: Financial Data Analysis - Retail and Telecommunication Industries - Science and Engineering - Intrusion Detection and Prevention - Data Mining and Recommender Systems - Data Mining and Society.

Text Books

1 Jiawei Han & Micheline Kamber, 2020, "Data Mining – Concepts and Techniques,",3rd Edition Elsevier

References

1 Ian H. Witten and Eibe Frank,2005, "Data Mining: Practical Machine Learning Tools and Techniques," 2nd Edition, Morgan Kaufmann

2 Arun K Pujari,2006, "Data Mining Techniques " 2ndEdition,University Press Publication

3 K.P Soman, Shyam Divakar, A.Ajay, 2006,"Insight into Data Mining: Theory and Practice ", 2nd Edition, Prentice Hall of India

4 Alex Berson, Stephen J. Smith, 2004, "Data Warehousing, Data Mining, & OLAP ", Tata McGrawHill

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2CC	INFORMATION AND NETWORK SECURITY	CORE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- The concepts of Information and Network security
- The issues regarding Confidentiality, Integrity, and Availability of a data
- The algorithms for Information and Network security

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Discuss the concepts of information security, threats, security services, and countermeasures.	K2
CO2	Illustrate the information security blueprints and major components	K3
CO3	Analyze various cryptographic algorithms	K4
CO4	Compare network security services and mechanisms	K4
CO5	Illustrate Internet security protocols for protecting data	K2

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		✓
CO2	✓	✓			✓
CO3	✓	✓	✓	✓	✓
CO4	✓	✓			
CO5	✓	✓			✓

COURSE FOCUSES ON

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2CC	INFORMATION AND NETWORK SECURITY	SEMESTER II

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Information Security 8 h

Information Security: Introduction - Concepts – Components of Information Systems – Approaches to Information Security Implementation – System Development Life Cycle – Security Systems Development Life Cycle – Security Professionals and Organization - Needs of Security: Business Needs – Threats – Attacks – Secure Software Development - Case Study : Theft of confidential Information

Unit II Security Planning and Technology 10 h

Information Security Planning and Governance – Information Security Policy, Standards and Practices – Information Security Blueprint - Security Technology : Access Control – Firewalls – Protecting Remote Connections – Intrusion, Detection and Prevention Systems – Honeypots, Honeynets and Padded Cell System.

Unit III Cryptography 10 h

Introduction: OSI Security Architecture – Security Services – Classical Encryption Techniques: Symmetric Cipher Model – Substitution Techniques – Transposition Techniques – Steganography –Block Cipher and Data Encryption Standards–Advanced Encryption Standard: AES Structure – AES Transformation Functions-RSA Case Study : Source Code Theft .

Unit IV Network Security 10 h

Network Access control – Extensible Authentication Protocol – Cloud Computing – Cloud Security risks and counter measures – Data protection in the cloud – Cloud security as a service - Addressing cloud computing security concerns – Transport Level Security: Web security Considerations - HTTPS - Case Study: Credit card Fraud.

Unit V Internet Security 10 h

Introduction: Internet Mail Architecture – Email Threats and comprehensive email security –Pretty Good Privacy –DNS Based Authentication of Named Entities – Sender Policy Framework –IP Security: Overview –Policy –Encapsulating security payload – Combining Security Associations – Internet key Exchange- Case Study : Phishing.

Text Books

1. Michael E. Whitman and Herbert J Mattord, 2011, "Principles of Information security", 4th Edition, Cengage Learning (Unit I,II)
2. William Stallings,2017,"Cryptography and Network Security Principles and Practices",7th Edition ,Prentice hall(Unit III-V)

References

1. Nina Godbole,2017,"Information Systems Security: Security Management, Metrics, Frameworks and Best Practices", 2nd Edition, Wiley
2. Micki Krause ,Harold F.Tipton,2008,"Handbook of Information security management", Vol 1-3, CRC press
3. Richard E. Smith, 2019, " Elementary Information Security ", 3rd Edition, Jones & Bartlett Learning
4. Jason Andress, 2019, "Foundations of Information Security: A Straightforward Introduction", No Starch Press (RHPS)

Course Code	Course Name	Category	L	T	P	Credit
235BA2A2EA	DIGITAL MARKETING ANALYTICS	EDC	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- The basic concepts of Digital marketing and its types, E-Business models.
- The importance of SEO and Digital Marketing Channels.
- To develop Digital technology in achieving marketing field.

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Explain the Digital marketing concepts, Challenges in E-Commerce Business	K1
CO2	Outline the Importance of SEO and Digital Marketing Channels	K2
CO3	Interpret the Strategies for Social Media Marketing and Analytics	K3
CO4	Understand the customization approaches for Digital Promotions	K2
CO5	Enumerate the Strategies and Services for Mobile and Web analytics	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓		✓	✓
CO2	✓	✓	✓		✓
CO3	✓	✓		✓	
CO4	✓	✓	✓	✓	✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development	✓	Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

235IB2A2EA	DIGITAL MARKETING ANALYTICS	SEMESTER II

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction to Digital Marketing 10 h

Digital Marketing – Origin and Development – Meaning -Strategy – Digital Marketing plan – Ethical and Legal Framework – Skills Required in Digital Marketing – Carriers in Digital Marketing – Digital vs. Real Marketing – Digital Marketing Trends and Challenges in E-Commerce and Business Models- Steps In Digital Marketing Plan- Planning Using Decision Trees.

Case Study on digital marketing strategy

Unit II Importance of SEO and Digital Marketing Channels 9 h

Search Engine Optimization (SEO) – The Concept of SEO – On-Page and Off-page Optimization -Search Engine Marketing - Blogs – Social Media Reach – E-mail Marketing – You Tube Marketing – Website Design and Hosting – Pay per Click and Search Engine Marketing- Metrics and Digital Marketing Channels.

Case Study on content marketing

Unit III Social Media Marketing 10 h

The role of social media marketing - Goals and strategies - Identifying target audiences - Social media platforms and social network sites - Micro blogging - Video marketing - Social media monitoring and maintenance- Social media marketing plan- Social Media Analytics

Case Study on social media marketing

Unit IV Types of Digital Promotions 8 h

Digital Marketing and Advertising – PPC advertising - Display Advertisement - Types of Digital Promotions – Print Advertising - Display Screens - Content Marketing and Customization Strategies – Digital Marketing and Understanding Buyer Behavior – Digital Advertising Market in India- Market Research - Market Information System - Market Survey.

Case Study on ethical challenges in digital marketing

Unit V Mobile Marketing and Web Analytics 11 h

Mobile advertising – Mobile marketing toolkit – Mobile marketing features – Mobile marketing on social networks- Mobile analytics- Web analytics – Key metrics – Making web analytics actionable –Types of tracking codes.

Case Study on Mobile Marketing

Note: Case Study examined externally (Section C: Compulsory question)

Text Books

1 Seema Gupta, 2020. Digital Marketing, 2nd Edition, McGraw Hill.

2 Puneet Bhatia, 2019. Fundamentals of Digital Marketing, 2nd Edition, Pearson Education.

References

1 Prashant Kadukar, 2020. The Power of Digital Marketing for Career & Business Success, 1st Edition, Digital Trainee Publications Pvt. Ltd.

2 Stephanie Diamond, 2010. Digital Marketing All - In - One for Dummies, 1st Edition, Wiley India Pvt Ltd.

3 Turban E, Lee J, King D, Liang T.P. and Turban D, 2015.Electronic Commerce, 8th Edition, Pearson Education.

4 Turban E,King D and Lee J,2015. Electronic Commerce: A Managerial and Social Networks Perspective,8th Edition,Prentice Hall Publisher.

234DA2A2CP	R FOR DATA ANALYTICS	**SEMESTER II**

Total Credits: 2
Total Instructions Hours: 48 h

S.No	List of Programs
1	Program to perform Vectors, Matrix Manipulation in R
2	Program to perform operations on Factors, Data Frame in R
3	Program for Operations on String and List in R
4	Program to Implement Regression using R
5	Program for Exploratory Data Analysis in R
6	Program to Import data sets and Perform read and write in R
7	Program to Implement Graphics in R
8	Program to Implement Cluster analysis in R
9	Program to Implement Time Series in R
10	Programs to implement Text preprocessing in R
11	Program to Implement Text Classification in R
12	Programs to implement Term frequency and Inverse Document Frequency in R

Note: Any 10 are Mandatory.

234DA2A2CQ	DATA MINING	SEMESTER II

Total Credits: 2
Total Instructions Hours: 48 h

S.No	List of Programs
1	Program to implement Linear regression for house price prediction
2	Program to implement SVM for .Sentiment analysis of reviews
3	Program to implement KNN for phishing website detection
4	Program to implement Bayesian classification for Credit card fraud detection
5	Program to implement Density based clustering
6	Program to implement Outlier detection
7	Program to implement customer segmentation
8	Program to implement Text mining
9	Program to implement Financial Data Analysis
10	Program to implement Multimedia data mining
11	Program to implement Movie recommender systems
12	Program to implement clickstream mining

Note: Any Ten Programs are mandatory.

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2DA	CUSTOMER ANALYTICS	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Concepts of Regression, Factor analysis, Demand forecasting and Market Basket Analysis
- Methods to increase Customer Lifetime Value and to identify Customer Loyalty
- Techniques in Predictive marketing to predict lead scoring and customer recommendations

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand regression and factor analysis	K2
CO2	Understand the techniques to forecast future demand	K2
CO3	Apply the concepts to increase customer lifetime value	K3
CO4	Analyze the methods in customer segmentation and profiling	k4
CO5	Apply the predictive marketing techniques for real world problems	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	✓
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓

COURSE FOCUSES ON

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2DA	CUSTOMER ANALYTICS	SEMESTER II

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Regression and Factor Analysis 9 h

Introduction to Retail Analytics – Sources of retail data - Regression – Assumptions – Need – Factor Analysis – Exploratory Vs confirmatory factor analysis – Using Factor Analysis - Understanding and estimating demand: Introduction - Regression to estimate demand - Properties of estimators - Time series data – Autocorrelation – Dummy variables – Business case

Unit II Forecasting Future Demand and Market Basket Analysis 10 h

Price elasticity and discounts – Introduction to elasticity – Modelling elasticity – Business case - Forecasting future demand: Autocorrelation – Dummy variables and seasonality – Business case - Product bundling: Market Basket - Logistic Regression - Predicting the market basket - Business case - Estimating time of Purchase: Conceptual overview of survival analysis – Business case - Model output and interpretation

Unit III Increasing Customer Lifetime Value and Identifying Loyalty 10 h

Increasing customer lifetime value: Descriptive analysis – Predictive analysis- Introduction to Tobit analysis - Business case – Modeling transactions - Quantifying complexity of customer behavior: Simultaneous equations – Need for Simultaneous equations - Missing value imputation - Identifying Loyal customers: Structural equation modeling – Business case

Unit IV Customer Segmentation and Profiling 10 h

Introduction to Segmentation: Segment – Strategic uses of segmentation – Apriori – Conceptual process – Tools for segmentation – Metrics- Chi Squared Automatic Interaction Detection (CHAID) – Drawing insights from segmentation – Business case - analytics - Introduction to Recency, frequency, monetary value (RFM) – Behavioral segmentation – RFM Vs Behavioral segmentation - Customer Profiling: Types of data to collect – Preparing data for analysis – Cleaning and validation – Linking and deduplication – Data integration

Unit V Predictive Marketing 9 h

Optimizing market spending using customer data - Predicting Customer Personas - Predicting the Customer Journey – Predicting the Customer Value and Value based marketing – Predicting Likelihood to buy - Choosing right level of discount using Likelihood to buy - Predictive lead scoring – Predicting Individual recommendations

Text Books

1. Mike Grigsby, 2016, "Advanced Customer Analytics - Targeting, Valuing, Segmenting and Loyalty Techniques" , Kogan Page Publishing
2. Omer Artun, Dominique Levin, 2015, "Predictive Marketing - Easy ways every Marketer can use Customer Analytics and Big Data", John Wiley & Sons

References

1. Mirza Rahim Baig, 2021, "Data Science for Marketing Analytics", Packt Publishing
2. Arvind Sathi, 2014, "Engaging Customers using Big Data", Macmillan Publishers
3. Stephanie Diamond, Marygrace Bateman, 2013, "Customer Analytics for Dummies", John Wiley & Sons
4. Wayne L. Winston, 2014, "Marketing Analytics: Data-Driven Techniques with Microsoft Excel", John Wiley & Sons

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2DB	NATURAL LANGUAGE PROCESSING	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Techniques of Natural Language Processing (NLP)
- Strategies for NLP system evaluation and error analysis
- Natural Language Generation (NLG) and Machine translation approaches

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the basic standard notation for characterizing text sequences	K2
CO2	Interpret word and sentence tokenization, and spelling error detection	K2
CO3	Evaluate N-gram language models by separating into training and test set	K4
CO4	Outline the concepts of semantic analysis	K3
CO5	Relate NLG applications and machine translation approaches	K4

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓	✓	
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓

COURSE FOCUSES ON

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2DB	NATURAL LANGUAGE PROCESSING	SEMESTER II

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction 10 h

Knowledge in Speech and Language Processing – Models and Algorithms – Language, Thought and Understanding - History – Natural Language Processing (NLP) Applications – Challenges of NLP – Regular Expressions and Automata: Regular Expressions – Finite-State Automata (FSA) – Regular Languages and FSAs

Unit II Words and Transducers 10 h

Finite - State Morphological Parsing – Construction of a Finite - State Lexicon -Finite - State Transducers - Sequential Transducers and Determinism - FSTs for Morphological Parsing - Transducers and Orthographic Rules - Combining FST Lexicon and Rules - Lexicon-Free FSTs - Word and Sentence Tokenization - Detecting and Correcting Spelling Errors - Minimum Edit Distance

Unit III Word Processing 8 h

N-Grams – Word Counting – Training and Test sets – Evaluating N-Grams – Smoothing – Parts of Speech (PoS) Tagging: English word classes – Tagsets – PoSTagging – Rule based POS tagging – HMM(Hidden Markov Model) PoS tagging –Transformation based tagging – Evaluation and error analysis

Unit IV Semantic Analysis and Discourse Processing 10 h

Semantic Analysis: Meaning Representation-Lexical Semantics - Ambiguity - Word Sense Disambiguation - Discourse Processing: cohesion - Reference Resolution - Discourse Coherence and Structure

Unit V Natural Language Generation (NLG) and Machine Translation 10 h

Architecture of NLG Systems – Generation Tasks and Representations – Applications of NLG – Machine Translation (MT): Problems in Machine Translation – Machine Translation Approaches – Direct – Rule-based – Corpus Based – Semantic or Knowledge-based MT systems

Text Books

1. Daniel Jurafsky, James H Martin, 2013, "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition", 2nd Edition, Pearson Education Inc.
2. Tanveer Siddiqui, U.S. Tiwary, 2008, "Natural Language Processing and Information Retrieval", Oxford University Press.

References

1. Masato Hagiwara , 2020, "Real World Natural Language Processing ", Manning
2. Steven Bird, Ewan Klein and Edward Loper, 2009, "Natural Language Processing with Python", 1st Edition, O'Reilly Media
3. James Allen, 1995, "Natural Language Understanding", Benjamin/Cummings Publishing Co.
4. Bikel, 2012, "Multilingual Natural Language Processing Applications: From Theory to Practice", 1st edition, Pearson Publishing

Course Code	Course Name	Category	L	T	P	Credit
234DA2A2DC	ADVANCED STATISTICS	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- The concepts of the Estimation theory
- The concept of design and Analysis of experiments
- The applications of non-parametric tests

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Interpret the Estimation theory	K2
CO2	Understand the different ways of testing hypothesis	K2
CO3	Analyze the test of Goodness of fit	K4
CO4	Compare analysis of experiment	K4
CO5	Apply non-parametric tests to practical problems	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓			
CO2	✓	✓		✓	
CO3	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	✓

COURSE FOCUSES ON

✓	Skill Development	✓	Entrepreneurial Development
✓	Employability	✓	Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A2DC	ADVANCED STATISTICS	SEMESTER II

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Estimation 10 h

Point Estimation: Introduction - unbiased Estimators - efficiency - consistency - sufficiency - robustness - the method of moments - the method of maximum likelihood - Bayesian estimation. Interval Estimation: Introduction - The estimation of means - the estimation of difference between means – the estimation of proportions - the estimation of difference between proportions – the estimation of variance – the estimation of the ratio of two variances – the theory in practice.

Unit II Hypothesis Testing 9 h

Introduction - testing a statistical hypothesis - losses and risks - The Neyman-Pearson lemma - the power function of a test - likelihood ratio test – the theory in practice

Unit III Tests of Hypothesis Involving Means, Variances and Proportions 9 h

Introduction - tests concerning means - tests concerning difference between means - tests concerning variances – tests concerning proportions - tests concerning differences among k-proportions - the analysis of an $r \times c$ table - goodness of fit – the theory in practice.

Unit IV Design and Analysis of Experiments 10 h

Introduction – one-way designs – randomized-block design - factorial experiments - multiple comparisons - other experimental designs – the theory in practice

Unit V Non Parametric Tests 10 h

Introduction - the Sign test - the signed-rank test - Rank-sum tests: The U Test - Rank-sum tests: the H test – tests based on runs - the rank correlation coefficient - the theory in practice

Text Books

1 Irwin Miller and Marylees Miller, John E. Freund's, 2007, "Mathematical Statistics with Applications", Seventh Edition, Prentices-Hall India Pvt Ltd, New Delhi.

References

1 Hogg and Craig, 2003, "Introduction to Mathematical Statistics", Pearson Education, New Delhi.

2 J.M. Kapur and H.C. Saxena, 2001, "Mathematical Statistics", S. Chand & Co, New Delhi.

3 Kandethody M. Ramachandran, Chris P. Tsokos, 2009, "Mathematical Statistics with Applications", Elsevier, Gurgaon.

4 Ronald E.Walpole, Raymond H. Myers, Sharon L. Myers, Keying E. Ye, 2018, "Probability and Statistics", Pearson Education, New Delhi.

16/10/23

BoS Chairman/HoD
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Science College
APPROVED

BoS- 9th	AC - 16th	GB - 21st
16.10.2023	13.12.2023	05.01.2024

Academic Council
CHAIRMAN
Dr.N.G.P. Arts and Science College

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3CA	MACHINE LEARNING	CORE	4	1	-	5

PREAMBLE

This course has been designed for students to learn and understand

- The foundation of machine learning and its practical applications
- The algorithms using training data to classify or predict the outcome of future datasets
- The evaluation of learning algorithms and model selection

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the Basic Concepts of Machine Learning.	K2
CO2	Apply algorithms and the paradigms of supervised Learning.	K3
CO3	Evaluate algorithms of Probabilistic Learning and Dimensionality Reduction.	K4
CO4	Apply algorithms and the paradigms of unsupervised learning.	K3
CO5	Implement various Reinforcement Learning algorithms.	K5

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓		✓	✓	✓
CO2	✓	✓			✓
CO3	✓	✓	✓	✓	
CO4	✓		✓	✓	✓
CO5	✓	✓	✓	✓	

✓	Skill Development		Entrepreneurial Development
✓	Employability	✓	Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3CA	MACHINE LEARNING	SEMESTER III

Total Credits: 5

Total Instruction Hours: 60 h

Syllabus

Unit I INTRODUCTION TO MACHINE LEARNING 12 h

Machine Learning – Types of Machine Learning – Supervised Learning – Unsupervised Learning – Reinforcement Learning – Machine Learning Process – Testing Machine Learning Algorithms – Model Evaluation and Improvement: Cross Validation - Grid search – A Brief Review of Probability Theory – The Bias-Variance Tradeoff

Unit II SUPERVISED LEARNING 12 h

Linear Regression - The Perceptron - Multi Layer Perceptron: Going Forwards – Backpropagation of Error - Support Vector Machines: Optimal Separation - Kernels - Tree based Models: Decision trees - Constructing Decision trees - ID3 – Classification and Regression Tree (CART) – Ensemble Learning: Boosting – Bagging – Case study

Unit III PROBABILISTIC LEARNING, DIMENSIONALITY REDUCTION 12 h

Probabilistic Learning: Naïve Bayes Classifier – Gaussian Mixture Models – Expectation Maximization Algorithm – Nearest Neighbor Methods: Nearest Neighbor Smoothing – KD Tree - Dimensionality Reduction: Linear Discriminant Analysis – Principal Component Analysis – Factor analysis – Independent Component Analysis – Locally Linear Embedding – Multi-Dimensional Scaling

Unit IV UNSUPERVISED LEARNING, OPTIMIZATION 12 h

Unsupervised learning: k-Means Algorithm – Dealing with Noise - Normalization – Weight Update Rule - Vector Quantization - Self-Organizing Feature Map - Genetic algorithms: Generating offspring - Using Genetic algorithms – Optimization: Going Downhill – Least – Squares optimization – Conjugate Gradients

Unit V REINFORCEMENT LEARNING, MARKOV CHAIN METHODS 12 h

Reinforcement Learning: Overview- State and Action Spaces- Reward function – discounting – Action Selection – Policy – Markov Decision Process- Values – Sarsa

and Q-learning – Markov Random Fields - Markov Chain Monte Carlo Methods: Sampling – Proposal Distribution - Markov Chain Monte Carlo – Hidden Markov Models - Viterbi Algorithm - Baum – Welch Algorithm - Case Study

Text Books

1. Ethem Alpaydın, 2015, "Introduction to Machine Learning", Third Edition, The MIT Press Cambridge, England,(Unit I,II).
2. Stephen Marsland, 2015, "Machine Learning: An Algorithmic Perspective", Second Edition, CRC Press,.(Unit III-V)]

References

1. Kevin P. Murphy, 2002, "Machine Learning A Probabilistic Perspective", MIT Press
2. Shai Shalev-Shwartz and Shai Ben-David, 2014, "Understanding Machine Learning from Theory to Algorithms", First Edition, University Press
3. C.M. Bishop, 2016, "Pattern Recognition and Machine Learning", Springer, Science Business Media, LLC
4. Stephen Marsland, 2014, "Machine Learning: An Algorithmic Perspective", Second Edition, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3CB	INTERNET OF THINGS AND APPLICATIONS	CORE	4		-	4

PREAMBLE

This course has been designed for students to learn and understand

- Smart Objects and IoT Architectures
- Various IoT related protocols
- Simple IoT Systems using Arduino and Raspberry Pi

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand IoT Network Architecture and Design	K2
CO2	Understand various IoT related protocols	K2
CO3	Build simple IoT systems using Arduino and Raspberry Pi	K5
CO4	Understand data analytics and cloud in the context of IoT	K2
CO5	Analyze IoT infrastructure for popular applications	K4

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓		✓	✓	✓
CO2	✓		✓	✓	✓
CO3	✓		✓		✓
CO4	✓	✓	✓		✓
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3CB	INTERNET OF THINGS AND APPLICATIONS	**SEMESTER III**

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction to IoT 9 h

Introduction: Genesis of IoT - IoT and Digitization - IoT Impact - Convergence of IT and IoT - IoT Challenges. IoT Network Architecture and Design: Drivers Behind New Network Architectures - Comparing IoT Architectures - A Simplified IoT Architecture - The Core IoT Functional Stack - IoT Data Management and Compute Stack

Unit II Engineering IoT Networks 10 h

Smart Objects : The Things in IoT : Sensors, Actuators and Smart Objects - Sensor Networks - Connecting Smart Objects : Communication Criteria - IoT Access Technologies. IP as the IoT Network Layer : The Business Case for IP - The Need for Optimization - Optimizating IP for IoT - Profiles and Compliances - Application Protocols for IoT : The Transport Layer - IoT Application Transport Methods : SCADA - CoAP - MQTT

Unit III Design and Development 10 h

Microcontroller Based Edge Devices: Microcontroller used in Ardunio - Microcontroller Peripherals - Microcontroller Programming Using Ardunio IDE - Physical Computing using Ardunio - Edge Device Signal Processing. Linux Based Edge Device Raspberry Pi: Raspberry Pi 3 Single Board Computer - Use Case of Smart Streetlight using Raspberry Pi Board - Interfacing and Programming Raspberry Pi Peripherals

Unit IV Data and Analytics for IoT 10 h

An Introduction to Data Analytics for IoT - Machine Learning - Big Data Analytics Tools and Technology - Edge Streaming Analytics - Network Analytics. Securing IoT : History of OT Security - Common Challenges in OT Security - IT and OT Security Practices and System Vary - Formal Risk Analysis Structures : OCTAVE and FAIR - The Phased Application of Security in an Operational Environment

Unit V IoT in Industry Applications 9 h

Manufacturing: Architecture for the Connected Factory - Connected Factory Security - Oil and Gas : Improving Opertional Efficiency - IoT Architectures for Oil and Gas - Smart and Connected Cities : Smart City IoT Architecture - Smart City Use - Case Examples - Transportation: IoT Use Cases for Transportation - Mining : - IoT Strategy for Mining - IoT Architecture for IoT in Mining - Public Safety: IoT Blueprint for Public Safety - IoT Public Safety Information Processing

Text Books

1. David Hanes, Gonzalo Salgueiro, Patrick Grossetete, Rob Barton and Jerome Henry, 2017, "IoT Fundamentals: Networking Technologies, Protocols and Use Cases for Internet of Things", Cisco Press (Units I - IV)

2. Rahul Dubey, 2019, "An Introduction to Internet of Things - Connecting Devices, Edge Gateway and Cloud with Applications", Cengage Learning India (Unit V)

References

1. Arshdeep Bahga, Vijay Madisetti, 2015, "Internet of Things – A hands-on approach", Universities Press

2. Olivier Hersent, David Boswarthick, Omar Elloumi, 2012, "The Internet of Things – Key applications and Protocols", Wiley

3. Jan Holler, VlasiosTsiatsis, Catherine Mulligan, Stamatis, Karnouskos, Stefan Avesand, David Boyle, 2014, "From Machine-to-Machine to the Internet of Things - Introduction to a New Age of Intelligence", Elsevier

4. Dieter Uckelmann, Mark Harrison, Michahelles, Florian (Eds), 2011, "Architecting the Internet of Things", Springer

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3CC	CLOUD COMPUTING	CORE	4		-	4

PREAMBLE

This course has been designed for students to learn and understand

- Different approaches for building the models and applications
- Complete Virtualization techniques
- Advanced Cloud services and platforms

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the basics of cloud computing.	K2
CO2	Understand various Fundamental Concepts and Models of Cloud	K2
CO3	Apply the Cloud Enabling Technology	K3
CO4	Apply Cloud Computing Architectures	K3
CO5	Understand Applications of Cloud Computing	K2

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓		✓	✓	✓
CO2	✓	✓			✓
CO3	✓	✓	✓	✓	
CO4	✓		✓	✓	✓
CO5	✓	✓	✓	✓	

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3CC	CLOUD COMPUTING	SEMESTER III

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Understanding Cloud Computing 8 h

Origin and Influences: History – Definitions – Business Drivers – Technology Innovations – Basic Concepts and Terminology: Cloud – IT Resource – On-Premise –Cloud Consumers and Cloud Providers – Scaling –Cloud Service – Cloud Service Consumer– Goals and Benefits –Risk and Challenges

Unit II Fundamental Concepts and Models 8 h

Roles and Boundaries – Cloud Characteristics – Cloud Delivery Models: Infrastructure-as-a-Service - Platform-as-a-Service - Comparing Cloud Delivery models – Combining Cloud Delivery models – Cloud Deployment Models: Public Clouds – Community Clouds – Private Clouds – Hybrid Clouds – Other Cloud Deployment Models

Unit III Basic Cloud Services, Platforms 12 h

Data Center Technology: Virtualization – Standardization and Modularity – Remote Operation and Management – High Availability – Secure Aware Design – Facilities – Computing Hardware – Storage Hardware – Network Hardware – Virtualization Technology: Hardware Independence – Server Consolidation – Resource Replication – Operating System Based Virtualization – Hardware Based Virtualization – Virtualization Management – Web Technology – Multitenant Technology – Service Technology

Unit IV Cloud Computing Architecture 12 h

Application Services - Content Delivery Services - Amazon CloudFront - Windows Fundamental Cloud Architecture : Workload Distribution Architecture – Resource Pooling Architecture – Dynamic Scalability Architecture - Service Load Balancing Architecture – Cloud Bursting Architecture – Advanced Cloud Architecture: Load Balanced Virtual Server Instances Architecture – Zero Downtime Architecture – Dynamic Failure Detection and Recovery Architecture – Fundamental Cloud Security: Cloud Security Threats

Unit V Cloud Platforms, Cloud Applications 8 h

Cloud Platforms in Industry: Amazon Web Services –Google App Engine – Cloud Applications: Scientific Applications – Healthcare: ECG Analysis in Cloud – Biology: Protein Structure Prediction – Business & Consumer Applications: CRM & ERP – Social Networking

Text Books

1. Thomas Erl, Zaigham Mahmood and Ricardo Puttini, 2019, "Cloud Computing Concepts, Technology and Architecture", Twelfth Impression, Pearson India Education Services Pvt. Ltd., Noida,. (Unit I to Unit IV)

2. Rajkumar Buyya, Christian Vecchiola and S.ThamaraiSelvi, 2018, "Mastering Cloud Computing",13th reprint, McGraw Hill Education(India) Pvt. Ltd, Noida,. (Unit V)

References

1. Michael Miller, 2008, "Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online", Kindle Edition, Que Publishing, Indianapolis US,.

2. Anthony TVelte, Toby JVelteand Robert Elsenpeter, 2010, "Cloud Computing –A Practical Approach", Tata McGraw Hill Education Pvt Ltd,.

3. Kumar Saurabh, 2011 , "Cloud Computing – Insights into New Era Infrastructure", 4th Edition., Pearson Education, Wiley Indian Edition, New Delhi,.

4. Kaittwang Geoffrey C.Fox and Jack J Dongrra, 2012, "Distributed and Cloud Computing", 1st Edition, Elsevier, USA,.

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3CD	BIG DATA ANALYTICS	CORE	4	1	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Different approaches for building the Hadoop MapReduce programs for real-time applications
- Complete Package of Big Data Analytics
- Mobile Analytics of Big Data on massive datasets

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand the structure of data and big data stack	K2
CO2	Understand the working environment of Hadoop ecosystem and its functionalities	K2
CO3	Apply Yarn, Pig and Hive to process big data	K3
CO4	Analyze Oozie, Casandra and Zookeeper and its uses for big data processing	K4
CO5	Apply Mobile Analytics for various real time applications	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓		✓	✓	✓
CO2	✓	✓			✓
CO3	✓	✓	✓	✓	
CO4	✓		✓	✓	✓
CO5	✓	✓	✓	✓	

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3CD	BIG DATA ANALYTICS	SEMESTER III

Total Credits: 4

Total Instruction Hours: 60 h

Syllabus

Unit I INTRODUTION TO BIG DATA 12 h

Big Data Introduction - Structuring big data - Elements of Big data - Big data Analytics - Future of Big data - Exploring the use of big data in business contexts : Preventive fraudulent activities - Retail industry - Understanding Big data technology foundation - Exploring the Big Data Stack - Ingestion layer - Virtualization and Big Data - Virtualization Approaches

Unit II HADOOP, MAPREDUCE 12 h

Understanding Hadoop Ecosystem : Hadoop Ecosystem - Hadoop Distributed File System - HDFS architecture - Command Line interface - Using HDFS file - Hadoop Specific File System Types - HDFS Commands – MapReduce - HBASE Architecture - Regions - Storing big data with HBASE - MapReduce framework - Working of MapReduce - Exploring Map and Reduce function

Unit III YARN, PIG, HIVE 12 h

Limitations of MapReduce - YARN architecture - Resource Manager - Application Manager - Working of YARN - YARN schedulers - Commands - User commands - Registry - Log Management in Hadoop - Hive: Introduction - Hive Services - Variables - Properties – Built-in functions in Hive - Data manipulation in Hive - Join Operations - Pig : Pig Architecture- Benefits - Operators - Debugging

Unit IV OOZIE, CASANDRA, ZOOKEEPER 12 h

Introduction of Oozie – Benefits - Oozie Workflow - Lifecycle Operations - Oozie parameterization : Workflow Functions - Coordinator Functions - Bundle Functions - Expression Language functions - Job Execution Model – Apache Cassandra : Features - Structure of database system - Read/Write Operations in Cassandra - Contextual Query Language - Zookeeper : Data Model - Znodes - Time, Session and Watches in Zookeeper

Unit V MOBILE ANALYTICS 12 h

Mobile analytics : Introduction - Types of results and Applications - Mobile analytics tools : Location based tracking tools - Real time analytics tools - User Behavior tracking tools - Performing mobile analytics - Challenges of mobile analytics

Applications : Online Social Media Analysis - Sentiment Analysis – Credit risk modeling - churn prediction-recommended systems - fraud detections - Business Process Analytics

Text Books

1 DT Editorial services, 2016, "Big Data Blackbook", First Edition, Dreamtech Press.

References

1 Michael Minelli, Michele Chambers,Ambiga Dhiraj, 2013, " Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business", First Edition, Wiley CIO Series.]

2 Tom White, 2012, "Hadoop: The Definitive Guide", Third Edition, O'Reilly Media.

3 Arvind Sathi, 2012, "Big Data Analytics: Disruptive Technologies for Changing the Game", First Edition, IBM Corporation,.

4 Bart Baesens, 2014, "Analytics in a Big Data World", Wiley Publishers.

234DA2A3CP	MACHINE LEARNING	SEMESTER III

Total Credits: 2
Total Instructions Hours: 48h

S.No	List of Programs
1	Programs to perform classification using Naive Bayes.
2	Programs to apply Linear Regression for house prediction dataset.
3	Programs to Implement SVM Classifier.
4	Programs to apply Decision tree algorithm for iris dataset
5	Programs to perform K means clustering.
6	Programs to apply K-Nearest Neighbor algorithm for healthcare dataset.
7	Programs to implement Logistic Regression.
8	Programs to implement Multilayer Perceptron.
9	Programs to implement Expectation Maximization algorithm.
10	Programs to implement Deep neural Networks.
11	Programs to implement Genetic algorithms
12	Programs to evaluate the performance of machine learning algorithms applied to a medical dataset

Note: Any 10 are Mandatory.

234DA2A3CQ	BIG DATA ANALYTICS AND VISUALIZATION	**SEMESTER III**

Total Credits: 2
Total Instructions Hours: 48 h

S.No	List of Programs
1	Perform the Extraction, Transformation, and Loading (ETL) process to construct the database using a data analytics tool.
2	Perform Imputation Methods in Data Preprocessing, grouping, and filtering using a data analytics tool.
3	Implement DAX (Data Analysis Expressions) queries using a data analytics tool.
4	Perform Geographical mapping on a massive dataset using a data analytics tool.
5	Create a Dashboard and Perform Data Analysis on given data set.
6	Program to perform Grouping and Joins Operations using Pig package.
7	Develop an application to find the maximum temperature using Spark.
8	Program for word count using MapReduce technique.
9	Program to create CRUD operations in Cassandra.
10	Program to import and export Cassandra data in HiveQL.
11	Program to create Views and indexes in HiveQL.
12	Program to sort and aggregate data in tables using HiveQL

Note: Ten Programs are mandatory

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3DA	DATA VISUALIZATION	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- To introduce visual perception and core skills for visual analysis.
- To Apply visualization techniques for various data analysis tasks.
- To understand visualization for multivariate analysis.

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Visualize the objects in different dimensions.	K1
CO2	Design and process the data for Virtualization.	K2
CO3	Apply the visualization techniques in physical sciences, computer science.	K3
CO4	Apply the visualization techniques in applied mathematics and medical sciences.	K3
CO5	Apply the virtualization Interaction Concepts.	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		✓
CO2		✓	✓	✓	✓
CO3	✓	✓		✓	✓
CO4	✓	✓			
CO5	✓	✓			✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3DA	DATA VISUALIZATION	**SEMESTER** III

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction and Data Foundation 8 h

Basics - Relationship between Visualization and Other Fields -The Visualization Process - Pseudo code Conventions - The Scatter plot. Data Foundation - Types of Data - Structure within and between Records - Data Preprocessing - Data Sets

Unit II Foundations For Visualization 10 h

Visualization stages - Semiology of Graphical Symbols - The Eight Visual Variables - Historical Perspective - Taxonomies - Experimental Semiotics based on Perception Gibson's Affordance theory – A Model of Perceptual Processing.

Unit III Visualization Techniques 10 h

Spatial Data: One-Dimensional Data - Two-Dimensional Data – Three Dimensional Data - Dynamic Data - Combining Techniques. Geospatial Data : Visualizing Spatial Data - Visualization of Point Data -Visualization of Line Data - Visualization of Area Data - Other Issues in Geospatial.

Unit IV Data Visualization Multivariate Data, Text and Document Visualization 10 h

Data Visualization Multivariate Data: Point-Based Techniques – Line Based Techniques - Region-Based Techniques - Combinations of Techniques – Trees Displaying Hierarchical Structures – Graphics and Networks- Displaying Arbitrary Graphs/Networks -Text and Document Visualization: Introduction - Levels of Text Representations - The Vector Space Model - Single Document Visualizations -Document Collection Visualizations.

Unit V Interaction Concepts 10 h

Extended Text Visualizations Interaction Concepts: Interaction Operators - Interaction Operands and Spaces - A Unified Framework. Interaction Techniques: Screen Space - Object-Space -Data Space -Attribute Space- Data Structure Space - Visualization Structure - Animating Transformations -Interaction Control

Text Books

1 Matthew Ward, Georges Grinstein and Daniel Keim,2010, "Interactive Data Visualization Foundations, Techniques, Applications ",A. K. Peters, Ltd,.

References

1 Alexandru C. Telea, 2008, "Data Visualization: Principles and Practice," A. K. Peters Ltd,.

2 Ben Fry,2008, "Visualizing data: Exploring and explaining data with the processing environment", O'Reilly,.

3 Scott Murray,2013, "Interactive data visualization for the web", O"Reilly Media, Inc.,

4 Robert Spence, 2007, "Information visualization – Design for interaction", 2 nd Edition, Pearson Education,

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3DB	MODERN DATABASES	DSE	4	-	-	4

PREAMBLE

This course has been designed for students to learn and understand

- Concepts of No-SQL Databases
- Types No-SQL databases
- Features of MongoDB, Neo4j

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Describe the features of No-SQL Databases	K1
CO2	Develop programs using Document and key value databases	K3
CO3	Experiment the features of column and XML data bases	K4
CO4	Apply the features of MongoDB for solving problems	K3
CO5	Construct queries and visualizations using graph databases and Neo4j	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1		✓		✓	
CO2	✓	✓	✓	✓	
CO3	✓	✓	✓	✓	
CO4	✓	✓	✓	✓	
CO5	✓	✓	✓	✓	

✓	Skill Development	✓	Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Social Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3DB	MODERN DATABASES	SEMESTER III

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction to NoSQL databases 08 h

Early Database Management Systems - Database revolutions: First, second and third generation - Big Data Revolution - Introduction to Sharding - Motivation for No-SQL Databases - CAP Theorem - Types of No-SQL Databases : Document Oriented - Columnar - Graph - Key - Value Pair- Case Study on Column Family Stores

Unit II Document Databases, XML databases 10 h

Introduction - Basic operation of document databases - XML and XML Databases: XML Tools and Standards - XML Databases - XML Support in Relational Systems - JSON Document Databases : Introduction - Data Models in Document Databases - Mem Base and Couch Base- Case Study on Key value stores

Unit III Column and Key-Value Databases 10 h

Introduction – Data Warehousing Schemas - The Columnar Alternative - Column Database Architectures - In-Memory Databases - Distributed Database Patterns: Distributed Relational Databases - Non-Relational Distributed Databases - Sharding and Replication- Case Study on XML Databases

Unit IV MongoDB and Database operations 10 h

Introduction to MongoDB: MongoDB Vs Relational Database Management Systems – Data Types – MongoDB Query Language - Getting Data into MongoDB– Database Operations: Create – Update – Read – Delete - Querying-Indexing - Aggregation - Comparison of Relational databases to new No-SQL stores - Case Study on MongoDB

Unit V Graph Databases, Neo4j 10 h

Graph databases: Need for graph databases - Neo4j : Key concept and characteristics -Modeling data for Neo4j : Four fundamental data constructs- Modeling for graph databases- Graph modeling - best practices - Importing data into Neo4j - Visualizations for Neo4j - Cypher Query Language - Case Study on Graph Databases

Text Books

1 Guy Harrison, 2015, "Next Generation Databases", 1st Edition, Apress

References

1 Shakuntala Gupta Edward, Navin Sabharwal, "Practical MongoDB",1st2 Edition,Apress.

2 Rik Van Bruggen, "Learning Neo4j", Second Edition, PacktPubishers, 2014

3 Adam Fowler, 2015, "NoSQL for Dummies", 1st Edition, John Wiley & Sons

4 Jeffrey A. Hoffer,V. Ramesh, 2016 , "Modern Database Management", First Edition, Pearson).

Course Code	Course Name	Category	L	T	P	Credit
234DA2A3DC	DEEP LEARNING	DSE	4		-	4

PREAMBLE

This course has been designed for students to learn and understand

- The fundamental techniques of Deep Neural Networks
- The architecture of Convolutional Neural Networks and Recurrent Neural Networks
- Transfer learning techniques and Natural language processing using Deep learning

COURSE OUTCOMES

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1	Understand Feed Forward Neural Networks	K2
CO2	Understand the concepts of Deep Neural Networks	K2
CO3	Apply Convolutional Neural Networks for various applications	K3
CO4	Analyze Recurrent Neural Networks and its types	K4
CO5	Apply Natural Language Processing and Transfer learning Techniques for real life problems	K3

MAPPING WITH PROGRAMME OUTCOMES

COs/POs	PO1	PO2	PO3	PO4	PO5
CO1	✓	✓	✓		✓
CO2		✓	✓	✓	✓
CO3	✓	✓		✓	✓
CO4	✓	✓			
CO5	✓	✓	✓	✓	✓

✓	Skill Development		Entrepreneurial Development
✓	Employability		Innovations
	Intellectual Property Rights		Gender Sensitization
	Awareness/ Environment		Constitutional Rights/ Human Values/ Ethics

234DA2A3DC	DEEP LEARNING	SEMESTER III

Total Credits: 4

Total Instruction Hours: 48 h

Syllabus

Unit I Introduction to Neural Network 10 h

Introduction to Neural Network - Model of Artificial Neuron - Learning rules and various activation functions Single layer Feed-forward networks -. Multi-layer Feed-forward networks - Recurrent Networks - Back Propagation networks - Architecture of Back-propagation Networks - Back propagation Learning

Unit II Deep Neural Networks 08 h

Introduction to Deep Neural Networks - Training deep models - Training Deep Neural Networks using Back Propagation - Setup and initialization issues - Gradient - Descent Strategies - Vanishing and Exploding Gradient problems – Regularization - Dropouts

Unit III Convolutional Neural Network 10 h

Basic structure of Convolutional Network - Convolutions for Images - Padding and Stride - Multiple Input and Multiple Output Channels – Pooling - Case study : Image classification, Object Detection, Image captioning using CNN

Unit IV Recurrent Neural Networks 10 h

Architectural Overview - Bidirectional RNNs – Encoder-decoder sequence to sequence architectures – Back-propagation Through Time for training RNN - Vanishing and Exploding Gradients - Long Short-Term Memory Networks - Gated recurrent Unit

Unit V Natural Language Processing, Transfer learning 10 h

Introduction to NLP - Word Vector representation - word2vec model - Continuous Skip Gram model - Continuous Bag-of-Words model. Case Study: Sentiment Analysis. Popular CNN Architectures and Transfer learning Techniques: LeNet, ResNet, VGGNet, AlexNet, DenseNet

Text Books

1 Ian Goodfellow, BengioY, and Courville A, 2016, "Deep Learning", MIT Press,

2 Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, 2021, "Dive into Deep Learning", Amazon Science

References

1 Hertz, John, Anders Krogh, and Richard G. Palmer, 1991, "Introduction to the Theory of Neural Computation", Redwood City, CA: Addison-Wesley Pub. Co,.

2 Simon S Haykin, 2006, "Neural Networks a Comprehensive Foundations" , PHIEd,.

3 Jacek M. Zurada, " Introduction to Artificial Neural Systems ", Jacek M. Zurada, JAICO Publishing House Ed,.

4 B. Vegnanarayana , 2005, "Artificial Neural Networks", Prentice Hall of India P Ltd,.

234DA2ASSA	BUSINESS ANALYTICS	SEMESTER III

Total Credits: 1

Total Instruction Hours:

Syllabus

Unit I Data Processing Architectures and BI

Key Purpose of Using IT in Business - Enterprise Applications - Types of Digital Data - OLTP - OLAP - OLAP Architectures - OLAP Operations - BI Component Framework - BI Users - BI Applications - BI Roles and Responsibilities

Unit II Data Warehouse and Data Mart

Need for Data Warehouse - Definition - Data Mart - ODS - Goals - Components of Data Warehouse - Extract, Transform, Load - Data Integration - Data Integration Technologies - Data Quality - Data Profiling.

Unit III Data Modeling Techniques

Data Modeling Basics - Types of Data Model - Data Modeling Techniques - Fact Table - Dimension Table - Dimension Models – Dimension Modeling Life Cycle - Enterprise Reporting - Balanced Scorecard - Dashboards

Unit IV Statistical Analysis

Statistics in Analysis - Data, Data Description and Summarization - Statistical Tests - Hypothesis and t-Test - Correlation Analysis - Regression - ANOVA - F-Test - Time Series Analysis

Unit V Applications of Analytics

Applications of Analytics in Healthcare - Applying Statistics to detect financial problems - Data Modeling in Advertising- Analytics for Help desk service improvement - Predictive Trendline Models - Forecasting Call Center Demand

Text Books

1 R N Prasad, Seema Acharya, Fundamentals of Business Analytics, 2016, 2nd Edition, Wiley India Pvt. Ltd

James R. Evans, Business Analytics, 2021, 3rd Edition, Pearson Education

References

1 S. Christian Albright, Wayne L. Winston, Business Analytics: Data Analysis & Decision Making, 2019, 6th Edition, Cengage Learning India.

2 Regi Mathew, Business Analytics for Decision Making, 2020, Pearson Education

Ramesh Sharda, Dursun Delen, Efraim Turban, Business Intelligence and Analytics: Systems for Decision Support, 2018, Pearson Education.

Asslani Arben, Business Analytics with Management Science Models and Methods, 2017, Pearson Education

234DA2ASSB	PROFESSIONAL ETHICS	SEMESTER III

Total Credits: 1

Total Instruction Hours:

Syllabus

Unit I Overview of Ethics

Ethics-Introduction –Importance of Integrity- Difference between Morals , Ethics and Laws- Ethics in the Business world- Corporate Social Responsibility - Creating an Ethical work Environment- Ethical Considerations in Decision Making - Ethics in Information Technology

Unit II Ethics for Information Technology Workers

IT Professionals - Professional Codes of Ethics- Professional Organizations– Licensing – Malpractice - IT Users - Common Ethical Issues For IT Users - Supporting Ethical Practices

Unit III Computer and Internet Crime

IT Security Incidents - Types of Exploits -Types of Perpetrators - Implementing Trustworthy Computing - Risk Assessment - Establishing a Security Policy - Educating Employees - Prevention - Detection - Response

Unit IV Privacy

Privacy Protection and the Law- Information Privacy - Privacy Laws-Applications--Key Privacy-Anonymity Issues-Data Breaches-Electronic Discovery-Consumer Profiling-Workplace Monitoring-Freedom of Expression

Unit V Intellectual Property

Introduction - Copyrights -Patents -Trade Secrets -Key Intellectual Property Issues - Plagiarism- Reverse Engineering - Open Source code-Competitive IntelligenceTrademark Infringement-Cybersquatting

Text Books

1 George W. Reynolds,(2014), Ethics in Information Technology, (5th Edn), Strayer University, Cengage Learning

References

1 Van de Poel, I., and L. Royakkers, (2011). Ethics, Technology and Engineering: An Introduction. Wiley-Blackwell..

2 R.S.Nagaarazan,(2006),"Professional Ethics and Human Values", (1st Edn.), New Age International Pvt. Limited.

BoS Chairman
Department of Computer Science with Data Analytics
Dr. N. G. P. Arts and Science College
Coimbatore – 641 048

Dr.N.G.P. Arts and Scien
APPROVED
BoS- 10th 2.4.24 | AC - 17th 17.4.24 | GB -

Academic Council
CHAIRMAN
Dr.N.G.P. Arts and Science College