MADHYA PRADESH BHOJ (OPEN) UNIVERSITY

PROGRAMME PROJECT REPORT

ON

BACHELOR OF SCIENCE (DATA SCIENCE)

Four Year Undergraduate Programme (Honours) offered as per NEP 2020

Submitted to

UNIVERSITY GRANTS COMMISSION

NEW DELHI

Submitted by

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY

BHOPAL (M.P.)

Contents

Madhya Pradesh Bhoj Open University, Bhopal

Programme Project Report

For

Bachelor of Science (Data Science) Four Year (Honours) Programme

1. Name of the Programme

BACHELOR OF SCIENCE (DATA SCIENCE) FOUR YEAR (HONOURS) PROGRAMME AS PER NEP 2020

2. Introduction

Bachelor of .Science (Data Science) Four-Year (Honours) programme is an u n d e r graduate programme that focuses on developing the skills and knowledge required to manage and analyze information technology in various industries. The program covers a wide range of topics such as software development, database management, computer networks, cybersecurity, and artificial intelligence.

Students who enroll in an Bachelor of .Science (Data Science) Four-Year (Honours) programme are typically individuals who already have an undergraduate degree in computer science, information technology, or a related field. The programme typically takes four years to complete, and students are required to complete coursework, research projects, and a dissertation.

The programme is designed to equip graduates with the necessary skills to become IT professionals, and prepares them for roles such as systems analyst, database administrator, network administrator, cyber security specialist, and software developer. This programme is highly valued in the job market, and graduates can expect to earn competitive salaries in their chosen field.

3. Programme's Mission:

The mission of Bachelor of Science .(Data Science) four-year (Honours) programme is to provide students with a comprehensive understanding of information technology and its applications in various industries. The programme aims to equip students with

the necessary skills to analyze, design, and develop IT systems that are efficient, effective, and secure.

The programme is designed to prepare students for careers in the rapidly evolving field of information technology, where there is a constant demand for professionals who can adapt to new technologies and solve complex problems. The programme emphasizes practical, hands-on experience, and encourages students to engage in research and development activities.

In addition, the programme aims to develop students' communication and teamwork skills, as these are essential for success in the field of IT. Graduates of the programme are expected to be leaders in their field, with the abilityto innovate, manage and lead IT projects.

Overall, the mission of this programme is to produce graduates who are well-rounded IT professionals, capable of contributing to the growth and success of organizations in various industries.

4. Objectives

The objectives of the Bachelor of Science (Data Science) four year (Honours) programme are as follows:

- **Developing advanced knowledge and skills:** The programme aims to provide students with advanced knowledge and skills in areas such as computer networks, database management, programming, artificial intelligence, cybersecurity, and more.
- **Preparing for leadership roles:** The programme often prepare students for leadership roles in the IT industry by providing them with the knowledge, skills, and confidence to lead teams and make strategic decisions.
- **Promoting innovation:** The programme encourage students to think creatively and innovatively by exposing them to cutting-edge technologies and research.
- **Enhancing career prospects:** The programme is designed to enhance students' career prospects by providing them with specialized knowledge and skills that are in high demand in the IT industry.

- **Encouraging lifelong learning:** The programme promotes a culture of lifelong learning by providing students with the tools and resources they need to stay up-to-date with the latest advancements inthe IT industry.

Overall, the objectives of this programme are to develop students' knowledge, skills, and competencies in science, preparing them for successful careers and lifelong learning, and enabling them to contribute to society through scientific inquiry and problem-solving.

5. Relevance of the Programme with HEI's Mission and Goals:

The relevance of B.Sc. (Data Science) four-year (Honours) programme is to fulfill institution's mission and goals which can vary depending on the specific institution. However, in general, this programme aligns with many higher education institutions' missions and goals in the following ways:

a. **Advancing scientific knowledge:** Many higher education institutions aim to advance scientific knowledge and research, and this programme plays an important role in achieving this goal. The programme provides students with a strong foundation in Data Science, enabling them to contribute to the advancement of scientific knowledge through research and discovery.

b. **Preparing students for careers in science:** Many higher education institutions aim to prepare students for successful careers, and this programme prepares students for careers in research, development, teaching, or any other Data Science-related profession.

c. **Fostering critical thinking and problem-solving skills:** Many higher education institutions aim to foster critical thinking and problem-solving skills, and this programme provides students with the opportunity to develop these skills through theoretical coursework, laboratory work, and research projects.

d. **Promoting interdisciplinary learning:** Many higher education institutions aim to promote interdisciplinary learning and collaboration, and this programme fosters interdisciplinary learning by enabling students to apply scientific principles and techniques to real-world problems in diverse fields.

e. **Encouraging lifelong learning:** Many higher education institutions aim to encourage lifelong learning, and this programme promotes lifelong learning by

fostering scientific inquiry and providing students with the skills and competencies required to continue learning throughout their careers.

f. **Meeting the needs of the IT industry:** Many HEIs aim to prepare their graduates to meet the needs of the industries they will work in. B.Sc.(Data Science) four year (Honours) programme do exactly that by equipping students with advanced knowledge and skills in areas such as computer networks, cybersecurity, and artificial intelligence, which are highly sought after by employers in the IT industry.

g. **Supporting research and innovation:** Many HEIs aim to support research and innovation by providing their students with access to cutting-edge technologies and research facilities. B.Sc.(Data Science) four year (Honours) programme often involves research projects, internships, and other opportunities for students to explore and develop innovative solutionsto real-world problems.

h. **Enhancing the institution's reputation:** Many HEIs aim to enhance their reputation by offering high-quality academic programs that produce successful graduates. B.Sc.(Data Science) four year (Honours) programme is highly regarded in the IT industry, and graduates of these programs are often sought after by employers, which can enhance theinstitution's reputation.

i. **Encouraging lifelong learning:** Many HEIs aim to promote a culture of lifelong learning by providing their students with the tools and resources they need to continue learning throughout their careers. B.Sc.(Data Science) four year (Honours) programme often involve ongoing professional development opportunities and access to industry experts, which can help students stay up-to-date with the latest advancements in the IT industry.

Overall, an B.Sc.(Data Science) program can be highly relevant to an HEI's mission and goals, particularly if the institution aims to prepare its graduates for successful careers in the IT industry, support research and innovation, and promote lifelong learning.

6. Nature of prospective target group of learners:

The prospective target group of learners for an B.Sc.(Data Science) four year (Honours) programme can vary depending on the specific program and institution offering it. However, in general, the target group of learners for an programme are individuals who:

- **Have an undergraduate degree in computer science, information technology, or a related field:** B.Sc.(Data Science) four year (Honours) programme typically requires applicants to have an undergraduate degree in a relevant field. This ensures that students have a solid foundation in IT fundamentals before embarking on more advanced coursework.
- **Are interested in advancing their knowledge and skills in IT:** B.Sc.(Data Science) four year (Honours) programme are designed for individuals who want to deepen their understanding of IT concepts, technologies, and applications. Prospective students may be interested in pursuing careers in areas such as software development, network administration, cybersecurity, or data analytics.
- **Are motivated and self-directed learners:** B.Sc.(Data Science) four year (Honours) programme require a high level of self- direction and motivation, as students must complete complex coursework, research projects, and other assignments with minimal supervision.
- **Have strong analytical and problem-solving skills:** B.Sc.(Data Science) four year (Honours) programme involve complex problem-solving and require students to think critically and analytically. Prospective students should have strong analytical and problem-solving skills to succeed in the programme.
- **Are interested in pursuing leadership roles in the IT industry:** B.Sc.(Data Science) four year (Honours) programme often prepares students for leadership roles in the IT industry by providing them with the knowledge, skills, and confidence to lead teams and make strategic decisions. Prospective students who are interested in pursuing leadership roles in the IT industry may find an programme particularly appealing.

Overall, the target group of learners for an B.Sc.(Data Science) four year (Honours) programme is typically composed of motivated, self-directed learners with a strong foundation in IT fundamentals who are interested in advancing their knowledgeand skills in IT and pursuing careers in the field.

7. Appropriateness of programme to be conducted in the Open and Distance Learning mode to acquire specific skills and competence:

B.Sc.(Data Science) four year (Honours) programme can be appropriately conducted in the Open and Distance Learning (ODL) mode to acquire specific skills and competence. Here are some reasons why:

- **Flexibility:** The ODL mode provides flexibility in terms of time, pace, and place of learning, which is especially important for learners who are already working or have other commitments. This mode allows learners to design their own study schedules and access course materials at their convenience, which can help them to balance their work, family, and educational commitments.
- **Access to learning resources:** ODL mode can provide access to a wide range of learning resources, including digital textbooks, audio and video lectures, interactive simulations, and virtual laboratories. This mode allows learners to access these resources from anywhere, which can help them to deepentheir understanding of key concepts and develop specific skills and competencies.
- **Personalized learning:** ODL mode can provide personalized learning experiences to learners by using adaptive learning technologies, providing individual feedback, and offering personalized tutoring. This mode can help learners to focus on their specific needs and interests and acquire the skills and competence they need to succeed in their chosen field.
- **Cost-effective:** ODL mode can be a cost-effective option for learners who are unable to attend traditional, on-campus programs. This mode can reduce the cost of tuition, accommodation, and transportation, making it more accessible to learners from diverse socioeconomic backgrounds.
- **Practical components:** Although some practical components of the program may require access to laboratory equipment, there are still many theoretical components of the program that can be delivered effectively through the ODL mode. The university offer online laboratories and simulations that can be used to supplement or replace the practical components.

Overall, the B.Sc.(Data Science) program can be appropriately conducted in the ODL mode, providing learners with the opportunity to acquire specific skills and competence in a flexible, accessible, personalized, and cost- effective manner.

8. Instructional Design

B.Sc. (Data Science) four-year (Honours) programme is having 160 credits. The course material for the programme has been developed in-house with contributions from seasoned academicians as well. The instructional design comprises all learning activities i.e., reading and comprehending the SLM, availing audio-visual aids to enhance knowledge, attending counselling sessions and preparing assignments.

The Programme is of four-year duration with annual examinations. The maximum period allowed is 8 years (double the duration). The Programme structure is as below.

(a) Detailed Syllabus of the Programme: Given as Annexure -01

(b) Duration of the Programme: The Duration of the Programme is four years.

(c) Faculty and Support Staff Requirement: At present, there are three faculty members in the Department. The Department have all support staff for its the smooth functioning.

(d) Instructional Delivery Mechanism:

The instructional delivery mechanisms of the programme should be designed to cater to the diverse needs of the learners, provide flexibility, and ensure effective learning outcomes. The programme should also provide adequate support services such as academic counselling, technical support, and mentorship to ensure that learners receive the necessary guidance and assistance throughout the programme.

As the University functions in the Open and Distance Learning mode, the programmes that we offer are designed to meet the varied requirements of the distant learner. Keeping this in view, the course material developed by the Department is learner friendly. Each course is divided into four to five blocks, which are further divided into units. Each Block consists of three to four units. This number has been determined taking into consideration the learning capabilities of our learners. The structure of the unit is in line with the guidelines laid down by the DEB-UGC, the apex regulatory body of Open and Distance Learning. The content is kept simple and lucid and follows the self- instructional pattern.

Each lesson includes a number of self-assessment questions along with hint answers so that the students are able to track their progress as they proceed with the lesson. At the end of each unit, a list of other relevant books is also provided. Besides providing quality study material to our learners, the Department, following the ODL pattern, has defined its programmes in the terms of credits. In the ODL system, one credit is equivalent to 30 study hours i.e. the study input required for completion of the programme. Normally the B.Sc (Data Science) four-year (Honours) programme is a 160-credit programme. This comprises all learning activities i.e., reading and comprehending the SLM, availing audiovisual aids to enhance knowledge, attending counselling sessions and preparing assignments.

9. Procedure for admission, curriculum transaction and evaluation

i. **Admission Process :**

Notification issued by the University in Regional and National Newspapers and in the official website. Admission process is online through the Online Portal. Payment of fee through online (various options like net banking etc.). Submission of the printout of the application by the candidate to concern study centre along with original documents for eligibility, date of birth etc., and along with fee paid receipt. After the Verification of applications - for fulfilment of eligibility criteria (marks cards) documents, fee paid details. Approval of the admission and issue of self-learning material (Study Materials) to the students.

ii. **Contact Programmes :**

The personal contact programme in every course shall extend over a period of 13 working days in each Year and is usually conducted at the beginning of the session. The students are expected to come prepared in the class in order to discuss their problems meaningfully. 75% attendance in the personal contact programme is mandatory.

iii. **Eligibility:** passed 12th standard in the Science stream.

iv. **Fee Structure:** B.Sc. Four Year Programme is Rs. 2960/- per year

v. **Evaluation norms:** A learner will be evaluated on the basis of Assignments and term-end examination. Assignments carry 30% weightage whereas the examination carry

70% weightage.

vi. Evaluation system:

- Students shall have a minimum of 50% of total marks of the University examinations in each Course. The overall passing minimum is 50% both in aggregate of Continuous Internal Assessment and External Examination in each Course.
- Every course shall have two components of assessment namely.

vii. Continuous Internal Assessment "CIA":

This assessment will be carried out throughout the Year as per the Academic Schedule. Continuous Internal Assessment for each Course shall be by means of Written Tests/ Assignments, and Class Tests for a total mark of 30. Continuous Internal Assessment for each Course shall be the responsibility of the concerned Course Faculty.

viii. Term-End Examination : This assessment will be carried out at the end of the Year as per the Academic Schedule.

ix. The valued answer papers/assignments shall be given to the students after the valuation is completed and they be asked to check and satisfy themselves about the marks they scored.

x. All records in respect of Continuous Internal Assessments shall be in the safe custody of the PI for at least one year after the assessment.

xi. Theory course assessment weightages:

The general guidelines for the assessment of Theory Courses, Department Electives and Non – Department Electives shall be done on a continuous basis as given in Table 1.

Table 1 : Weightage for Assessment

S.No.	Assessment	Weightage	Duration
1.	First Periodical Assessment	10%	2 periods

2.	Second Periodical Assessment	10%	2 Periods
3.	Practical/Project/ Lab	10%	--
4.	Term End Exam	70%	2 to 3 hours

xii. Grading System

Based on the student's performance in each Year, grade is awarded with a final letter grade at the end of the exam evaluation of each Course. The letter grades and the corresponding grade points are as follows.

Table 2: Grading system

Range of Marks	Letter Grade	Grade Points	Remarks
90 – 100	S	10	Outstanding
80-89	A	09	Excellent
70-79	B	08	Very Good
60-69	C	07	Good
50-59	D	06	Average
40-49	E	05	Pass
<40	U	00	To Reappear for Term-End Examination

xiii. GPA and CGPA

Grade Point Average (GPA) is the ratio of the sum of the product of the number of credits C_i of course "i "and the grade points Pi earned for that course taken over all courses "i" registered and successfully completed by the student to the sum of C_i for all "i". That is,

$$GPA = \frac{\sum_1^n \; C_i \; P_i}{\sum_1^n \; C_i}$$

Cumulative Grade Point Average (CGPA) will be calculated in a similar manner, in any Year, considering all the courses enrolled from the first Year onwards. The Grade card will not include the computation of GPA and CGPA for courses with letter grade "U" until those grades are converted to the regular grades.

xiv. Grade Sheet

Based on the performance, each student is awarded a final letter grade at the end of the Year in each course. The letter grades and corresponding grade points are given in Table 2. A student is considered to have completed a course successfully and earned credits if he/she secures a letter grade other than U in that course. After results are declared, grade sheet will be issued to each student which will contain the following details:

xv. Programme and discipline for which the student has enrolled.

xvi. Year of registration.

xvii. The course code, name of the course, category of course and the credits for each course registered in that Year

xviii. The letter grade obtained in each course

xix. Year Grade Point Average (GPA)

xx. The total number of credits earned by the student up to the end of that Year in each of the course categories.

xxi. The Cumulative Grade Point Average (CGPA) of all the courses taken from the first Year.

xxii. Credits earned under Non – CGPA courses.

xxiii. Additional credits earned for the respective UG degree or respective UG degree with Minor specialization.

xxiv. Class/Division

Classification is based on as follows: CGPA ≥ 8.0: First Class with Distinction

6.5 ≤ CGPA <8.0: First Class

5.0 ≤ CGPA <6.5: Second Class

(i) Further, the award of 'First class with distinction' is subject to the candidate becoming eligible for the award of the degree, having passed the examination in all the courses in his/her first appearance with effect from first-year and within the minimum duration of the programme.

(ii) The award of 'First Class' is further subject to the candidate becoming eligible

for the award of the degree, having passed the examination in all the courses within 8 years.

(iii) The period of authorized break in study will not be counted for the purpose of the above classification.

xxv. Eligibility For the Award of Degree

A student will be declared to be eligible for the award of the B.Sc. four-year (Honours) programme if he/she has registered and successfully obtained credit for all the core courses:

i. Successfully acquired the credits in the different categories as specified in the curriculum corresponding to the discipline of his/her study within the stipulated time:

ii. Has no dues to all sections of the institute including Guest house/hostels and has no disciplinary action pending against him/her.

iii. The award of the degree must be recommended by the Academic Council and approved by the Board of Management of the university.

xxvi. Re-View Of Answer Scripts / Single Valuation

If any student feels aggrieved on the final outcome of the assessment in any course, the student shall apply to the Controller of Examinations, along with the prescribed fee, for the review of Term End examination answer script, within the stipulated time after the announcement of the results of the examinations. The Controller of Examinations shall facilitate the review of the answer script jointly to be carried out by the student and the faculty detailed for this purpose. If any discrepancy is noticed during review the same shall be rectified and the originally awarded grade accordingly amended.

10. Requirement of the laboratory support and library resources:

i. Internet Leased Line

ii. Computer Lab:

The University has most modern high-tech Computer Lab with 24 hrs. Internet facility for

studying and R&D activities. The state-of-the-art facility features hardware & software that is geared to specific academic programmes.

iii. Research Lab:

University equipped with round the clock available Hi-tech research Lab. which includes latest configured systems with Hi-speed internet facility loaded with latest software for research purpose.

iv. Library:

The University library is the hub of knowledge with more than 105832 books, Online Journals/Magazines, Back Volumes, Thesis & Dissertations:, News Papers: 12, Book Bank: 25850 are conveniently accessible for the students and staff of MPBOU. Library includes Main reading area, separate reading area and reference section.

The MPBOU system consists of a Central Library and Departmental Libraries which collectively support the teaching, research and extension programmes of the Institute. All students, faculty members and employees of the Institute are entitled to use the Library facilities enrolling membership.

v. Internet Facilities:

One Gbps – Lease Line Link are available at the university. University is fully networked with a campus wide network interconnecting all departments. Campus is fully Wi-Fi with high-Speed internet connection available round-d-clock.

vi. MS TEAMS Platform:

The University has acquired the Zoom platform license and has gone 100% virtual in terms of delivering the regular classes during the COVID lockdown. This has given enough experience for the faculty and expertise to handle the platform for online learning. More than 150 guest lecturers have been organised in the brief period of three weeks inviting industry experts. The MPBOU team is now digitally empowered to conduct OL classes on this virtual platform on a regular basis which shall make the learning process very effective.

LMS: The Institute is using MOODLE and Microsoft Teams for Online teaching and to

conduct test and Assignments.

11. Cost estimate of the Programme and the provisions:

The B.Sc. four-year (Honours) programme in basic sciences has been prepared in-house. The internal faculty has made contribution through units. The Department has utilized the acumen and expertise of seasoned academicians in the development of the course material. The expenses incurred in outsourcing units is as follows:

Cost Estimates for Development of the Programme:

Programme	Programme Development	Delivery	Total
B.Sc. (Data Science)	7,00000/-	3,00000/-	10,00000/-

12. Quality Assurance Mechanism and Expected Programme Outcomes

The Quality Assurance Mechanism and expected programme outcomes for the B.Sc. four-year (Honours) programme are as follows:

i. Quality Assurance Mechanism:

a. **Curriculum Design:** The programme curriculum is designed in consultation with industry experts, academics, and professional bodies to ensure that it is relevant and up-to-date.

b. **Course Delivery:** The course delivery is supported by the latest technologies, experienced faculty, and a range of learning resources such as virtual labs, simulations, and online discussion forums.

c. **Assessment:** The assessment system is rigorous, transparent, and fair, with regular formative and summative assessments, which measure the students' learning outcomes.

d. **Feedback:** The feedback system is designed to ensure that students receive timely and constructive feedback on their performance, enabling them to identify areas of improvement and take corrective action.

e. **Quality Control:** The quality control mechanisms include regular internal and external reviews and evaluations, ensuring that the programme's quality is maintained and improved continuously.

ii. Expected Programme Outcomes:

f. **Scientific knowledge:** Students will acquire a strong foundation in scientific principles and concepts, enabling them to apply their knowledge in diverse fields.

g. **Critical thinking:** Students will develop critical thinking skills, enabling them to analyze and solve complex problems.

h. **Communication:** Students will develop effective communication skills, enabling them to communicate scientific ideas and findings to a wide range of audiences.

i. **Research and experimentation:** Students will develop research and experimentation skills, enabling them to conduct scientific research and experimentation.

j. **Professionalism:** Students will develop professional and ethical values, enabling them to work effectively in a team and contribute to society.

Overall, the Quality Assurance Mechanism and expected programme outcomes for the B.Sc. (Data Science) four-year (Honours) programme ensure that students receive a high-quality education, which prepares them for successful careers in research, development, teaching, or any other science-related profession.

Fieldwork/Internship/Project

Guidelines

A. **PROJECT FORMULATION**:

This component of Fieldwork/Internship/ Project may be taken with the following perspectives

- The component should be done in an application area of admitted programme having relevant field.
- Comprehensive case study (covering single organization/multifunctional area problem,formulation analysis and recommendations) also may be conducted.
- Inter-organisational study aimed at inter-organisational comparison also may beconducted.
- Evolution of any new conceptual / theoretical/practical framework.
- Field study (empirical study).
- It may be based on primary or secondary data.

B. PROPOSAL/SYNOPSIS/PLAN OF ACTION:

Learners will have to submit when the detailed guidelines will be made just before commencement.

C. TYPE OF THE FIELDWORK/INTERNSHIP/PROJECT:

The learners are expected to work on the scientific idea related to the relevant topic/interdisciplinary field of the programme preferably in some specific issue/topic related to selected course of the admitted programme. Learners are encouraged to work in the areas closely associated with their programme of study. The learner can formulate a project problem with the help of her/his Guide and submit the project proposal for approval to the Coordinator of the respective study centres. Approval of the project proposal is mandatory. If approved, the learner can commence working on it, and complete it.

D. INTERNSHIP

i. The learner if not employed elsewhere, may choose an organization for internship projects. The University will help the learner in conceptualizing the proposal in consultation with the organizational guide.

ii. The learner will have to submit the Internship Report with an endorsement certificate from the organisation.

iii. The learner will be required to submit a weekly diary of the activities conductedduring the Internship period.

E. STEPS INVOLVED IN THE PROJECT WORK:

The project work should be done by the learner only. The role of the guide should be about guidance wherever any problem encounters during the preparation of project. The followingare the major steps involved in the preparation of project, which may help you to determinethe milestones and regulate the scheduling of the project:

- Select a guide in consultation with the coordinator (be in touch with the guide during the work).
- Select a topic.
- Prepare the project proposal in consultation with the project guide.
- Submit the project proposal (two copies, one to be returned to the candidate afterapproval) along with the necessary documents to the Coordinator of the study centre.
- Receipt of the project approval from the Coordinator of the study centre.
- Carry out the project-work.
- Prepare the project report.
- Submit the project report to the Coordinator of the study centre within 4 months. from the last date of last examination of 3rd Semester.
- Appear for the viva-voce as per the schedule declared by the University.

F. RESUBMISSION OF THE PROJECT PROPOSAL IN CASE OF NON-APPROVAL:

i. In case of non-approval, the suggestions for reformulating the project will be communicated to the learner by the Centre Coordinator. The revised project synopsis along with a new proforma, should be re-submitted along with a copy

of the earlier synopsis and non-approval project proposal proforma to the Centre Coordinator.

ii. If the learner wants to undertake a new project by changing his/her earlier project proposal, he will have to justify his new choice. Without valid ground and certification from his/her guide, no change in project proposal will be entertained.

iii. In any case, changes in project proposal will not be allowed after submitting the second project proposal. The second proposal will be considered as final.

iv. It is necessary that the learners finalize their project proposal well ahead of time.

v. It is to be noted that changes in project proposal will not be entertained in the last month of the programme.

vi. In order to complete the project in due time, a learner should devote at least 120 days for his/her project. This time should be judiciously divided into various phases like field study & interview, data collection, data tabulation, data interpretation and data analysis if the project is based on fieldwork.

vii. If the learner chooses an academic area concerning evolution of any new conceptual / theoretical/practical framework, a synopsis needs to be submitted with adequate review of literature, and formulation of research objectives, research questions and hypotheses asmay be applicable.

G. PROJECT PROPOSAL FORMULATION:

The project proposal should be prepared in consultation with the guide. The project proposal should clearly state the project objectives. The project proposal should contain complete details in the following form:

- Proforma for Approval of Project Proposal duly filled and signed by both the learner and the Project Guide with date.
- Synopsis of the project proposal (4-6 pages) covering the following aspects:
  - (i) Title of the Project
  - (ii) Introduction and Objectives of the Project
  - (iii) Methodology
  - (iv) Project Planning and Scheduling
  - (v) Reference.

- Violation of the project guidelines will lead to the rejection of the project at any

stage.

H. PROJECT REPORT SUBMISSION:

After completion of the work two copies of the report need to be submitted to the Centre Coordinator and a copy of the report should be kept by the candidate. The candidate should carry the copy of the report at the time of viva voce examination and get it verified and signed by the examiner. Out of the two copies submitted to the respective Study Centre Coordinator, one copy is to be marked as 'University Copy' and the second copy is to be marked as 'Study Centre Copy'.

I. ASSESSMENT GUIDELINES FOR PROJECT EVALUATION:

Each component of the project work and the viva voce carries its own weightage, so thelearner needs to concentrate on all the sections given in the project report formulation.

J. PROJECT EVALUATION:

The Project Report is evaluated for 200 marks. Viva-voce is compulsory and forms part of evaluation. A learner in order to be declared successful in the project must secure 40% marksin each component (i) Project Evaluation and (ii) Viva-voce. The learners must compulsorily clear both the components of the project. If a learner submitted the project report as per the schedule and fails to attend viva, her/his Project will remain incomplete and should contactthe Coordinator of the study centre. The candidate may use power point for viva-voce in consultation with the project guide.

K. ARRANGEMENT OF CONTENTS:

The sequence in which the project report should be arranged and bound should be as follows:

1. Cover Page
2. Inside cover page
3. Certificate of Originality by the Project Guide
4. Acknowledgement
5. Abstract
6. Table of Contents

7. List of Tables
8. List of Figures
9. Chapters
10. References
11. PAGE DIMENSION AND BINDING SPECIFICATIONS: The project report should beprepared in A4 size, and should be in spiral binding. Abstract –Abstract should be a one page synopsis of the project work, typed in doubleline spacing (about 300 words with maximum 6 key words). Font Style Times New Roman and Font Size 12.
12. Table of Contents–The table of contents should list all headings and sub-headings. The title page and Certificate will not find a place among the items listed in the Table of Contents. 1.5 line spacing should be adopted for typing the matter under this head.
13. List of Tables–The list should use exactly the same captions as they appear above thetables in the body of the report. 1.5 line spacing should be adopted for typing the matter under this head.
14. List of Figures –The list should use exactly the same captions as they appear below the figures in the body of the text. 1.5 line spacing should be adopted for typing the matter under this head.
15. Chapters- The chapters may broadly be divided into the following parts with minimum total number of pages ranging from 50 to 700-
    (i) Introductory chapter
    (ii) Objectives and importance (significance) for the Study,
    (iii) Research Methodology opted for the Study (stating nature, sources, collection ofdata; research tools and techniques to be used; sampling procedure)
    (iv) Limitations and Scope of the Study
    (v) Data Analayisis and Interpretaion
    (vi) Conclusion
    (vii) References
16. List of References- The listing of references should be typed 4 spaces below the heading “REFERENCES” in alphabetical order, in single spacing and left-justified. Typing Instructions- The impression on the typed copies should be black in colour. 1.5 line spacing should be used for typing the general text. The

general text shall be typed in the Font style: 'Times New Roman' and Font size: 12. Use A4 (210 mm X 297 mm) for all copies submitted. Use only one side of the paper for all printed/typed matter.

Page Numbering- Every page in the project report, except the project report title page, must be numbered. The page numbering, starting from acknowledgement and till the beginning of the Introductory chapter should be printed in small Roman numbers, i.e, i, ii, iii, iv..... The page number of the first page of each chapter should not be printed. Allpage numbers from the second page of each chapter should be printed using numerical,

i.e. 2,3,4,5... All printed page numbers should be located at the bottom centre of the page. Chapter Numbering- Use only Roman numerical. Chapter numbering should be centered on the top of the page using large bold print. < size 15> TEXT-COLOR -Black SECTION HEADINGS -Times New Roman 12 pts., bold print and all capitals. SUBSECTION HEADINGS -Times New Roman 12 pts., bold print and leading. capitals. ie. Only first letter in each word should be in capital. JUSTIFICATION The text should be fully justified MARGINS The margins for the regular text are as follows:

LEFT = 1.50"

RIGHT = 1.00"

TOP = 1.00"

BOTTOM = 1.00

L. ELIGIBILITY CRITERIA OF A PROJECT GUIDE

i. A Faculty of the relevant School having degree in relevant field or area (subject matter expert) , MPBOU headquarter.

ii. Any university teachers in the relevant scientific field or area with minimum five years of teaching experience (Preferably PhD /M.Phil).

iii. Any college teachers in the in relevant field or area (subject matter expert) affiliated to any Indian University recognized by UGC and having minimum

five years of teaching experience. (Preferably PhD /M.Phil).

iv. Faculties in the departments of relevant School having degree in relevant field or area (subject matter expert) which are recognized Study centres of MPBOU and having minimum five years of teaching experience.

M. ASSESSMENT OF SEMINAR/PRESENTATION/PROJECT AND DISSERTATION

A learner has to make seminar/presentation and prepare a project/dissertation on a particular topic under the guidance of a professionally qualified supervisor/guide in Semester VII and VIII of proposed Programme. They should devise instruments for the collection and interpretation of data and the preparation of research report. The length of the report may be between 60-70 pages.

Academic Structure : UG Program

		own Faculty		Any Faculty	skill Enhancement course	Ability Enhancement course	Field Project/internship/ apprenticeship/community engagement and service	credits		Qualification title (credit requirement) Entry/Exit Criteria
		subject I	subject II	subject III						
Eligibility will have pre-requisites				Elective course	vocational course	Foundation course	Inter/Intra Faculty related to main subject	credits Distribution	Total Credits per year	
year		Major	Minor							
LEVEL 5	1	2 (12 credits) 6 credits each	1 (06 Credits)	1 (06 Credits)	1 (4 Credits)	2 (8 Credits) 4 Credits each	1 (4 Credits)	12+6+6 4+8+4	40	(40) Undregraduate certificate in Faculty
LEVEL 6	2	2 (12 credits) 6 credits each	1 (06 Credits)	1 (06 Credits)	1 (4 Credits)	2 (8 Credits) 4 Credits each	1 (4 Credits)	12+6+6 4+8+4	40	(80) Undregraduate Diploma in Faculty
LEVEL 7	3	2 (12 credits) 6 credits each DSE	1 (06 Credits)	1 (06 Credits)	1 (4 Credits)	2 (8 Credits) 4 Credits each	1 (4 Credits)	12+6+6 4+8+4	40	(120) Bechelor in Faculty
LEVEL8	4	2 (12 credits) 6 credits each 2 (08 credits) 4 credits each	1 Research methodology (4 credits) 1 (4credits)				1 (12 Credits) (6+6) internship/appenticeship related to main subject/Research Project	20+4+4+12	40	(160) Bechelor (honours/Research) in Faculty
				Credits are given in (in red)						
Total		56 credits	26 Credits	18 credits	12 credits	24 credits	24 credits		160 credits	

Bachelor of Science in Data Science

(B. Sc. Data Science)

First Year

S. No	Course Code	Subject	Max Marks		Total Marks		Total Credits
Major			**Assign**	**Theory**	**Max**	**Min**	
1		**Course- 1:** Computer Fundamentals, Organization and Architecture	30	70	100	35	4
2		**Course-2:** Programming Methodology And Data Structure	30	70	100	35	4
Minor							
3		Operating System	30	70	100	35	4
Elective							
4		Introduction to Data Science & Python Programming	30	70	100	35	4
Vocational							
5		Web Designing	30	70	100	35	2
LAB							
6		Computer Fundamentals, Organization and Architecture Lab			100	35	2
7		Programming Methodology And Data Structure Lab			100	35	2
8		Operating System Lab			100	35	2
9		Introduction to Data Science & Python Programming Lab			100	35	2
10		Web Designing Lab			100	35	2

Second Year

S. No	Course Code	Subject	Max Marks		Total Marks		Total Credits
Major			**Assign**	**Theory**	**Max**	**Min**	
1		**Course- 1** DBMS and Data Mining	30	70	100	35	4
2		**Course-2** Big Data and Big Data Analysis	30	70	100	35	4
Minor							
3		Artificial Intelligence and Machine Learning	30	70	100	35	4
Elective							
4		Advance web technology	30	70	100	35	4
Vocational							
5		Introduction to Data Science and R Programming	30	70	100	35	2
LAB							
6		DBMS and Data Mining Lab			100	35	2
7		Big Data and Big Data Analysis Lab			100	35	2
8		Artificial Intelligence and Machine Learning Lab			100	35	2
9		Advance web technology Lab			100	35	2
10		Introduction to Data Science and R Programming Lab			100	35	2

Third Year

S. No	Course Code	Subject	Max Marks		Total Marks		Total Credits
Major			**Assign**	**Theory**	**Max**	**Min**	
1		**Course- 1:** Data Mining Concepts and Techniques ta Mining Concepts and Techniques	30	70	100	35	4
2		**Course- 2** Python Programming For Data Analysis	30	70	100	35	4
Minor							
3		Big Data Analytics Using Spark	30	70	100	35	4
Elective							
4		Data Visualization	30	70	100	35	4
Vocational							
5		Data Analytics With Tableau	30	70	100	35	2
LAB							
6		Data Mining Concepts and Techniques ta Mining Concepts and Techniques Lab			100	35	2
7		Python Programming For Data Analysis Lab			100	35	2
8		Big Data Analytics Using Spark Lab			100	35	2
9		Data Visualization Lab			100	35	2
10		Data Analytics With Tableau Lab			100	35	2

Fourth Year

S. No.	Course Code	Subject	Max Marks		Total Marks		Total Credits
Major			**Assign**	**Theory**	**Max**	**Min**	
1		Course- 1 AI Concepts and Techniques With Python	30	70	100	35	4
2		Course-2: Supervised Ml With Python					
Minor							
3		Unsupervised Ml With Python	30	70	100	35	4
Elective							
4		NLP With Python	30	70	100	35	4
		Vocational					
5		Deep Learning Neural Networks With Python	**30**	**70**	**100**	**35**	**2**
LAB							
6		AI Concepts and Techniques With Python Lab			100	35	2
7		Supervised Ml With Python Lab			100	35	2
8		Unsupervised Ml With Python Lab			100	35	2
9		NLP With Python Lab			100	35	2
10		Project			100	35	2

Bachelor of Science in Data Science

(B. Sc. Data Science – First Year)

Computer Fundamentals, Organization and Architecture

Unit 1

Fundamentals of computers: Definition, Characteristics, capabilities and limitations . Types of Computers: Analog, Digital, Micro, Mini, Mainframe & Super Computers, Work Station, Server computers. Generations of Computers. Smart Systems: definition, characteristics and applications. Definition of Embedded system, GIS, GPS, Cloud Computing. Uses of computers in e-governance and various public domains and services.

Unit II

Block diagram of computer and its functional units. Concept of Hardware, software and firmware. Types of software. Input devices - keyboard, scanner, mouse, light pen, bar code reader, OMR, OCR, MICR, track ball, joystick, touch screen camera, microphone etc. Output devices: monitors classification of monitors based on technology -CRT & flat panel, LCD, LED monitors, speakers, printers - dot matrix printer, ink jet printer, laser printer, 3D Printers, Wi-Fi enabled printers, plotters and their types, LCD/LED projectors. Computer memory and its types, Storage devices: Magnetic tapes, Floppy Disks, Hard Disks, Compact Disc - CD-ROM, CD-RW, VCD, DVD, DVD-RW, usb drives, Blue Ray Disc, SD/MMC Memory cards .

Unit III

Fundamentals of Digital Electronics: Data Types, Complements, Fixed-Point Representation, Floating-Point Representation, Binary and other Codes, Error Detection Codes. Logic Gates, Boolean Algebra, Map Simplification, Combinational circuits, Sequential Circuits, simple combinational circuit design problems. Combinational Circuits- Adder- Subtractor, Multiplexer, Demultiplexer, Decoders, Encoders Sequential Circuits - Flip-Flops, Registers, Counters.

Unit IV

Basic Computer Organization: Instruction codes, Computer Registers, Computer Instructions, Timing & Control, Instruction Cycles, Memory Reference Instruction,
Input - Output & Interrupts Instruction formats, Addressing modes, Instruction codes, Machine language, Assembly language.Register Transfer and Micro operations: Register Transfer Language, Register Transfer, Bus & Memory Transfer, Arithmetic Micro- operations, Logic Micro-operations, Shift Micro-operations.

Unit V

Processor and Control Unit: Hardwired vs. Micro programmed Control Unit, General Register Organization, Stack Organization, Instruction Format, Data Transfer & Manipulation, Program Control, Introductory concept of RISC, CISC, advantagesandisadvantages of both. Pipelining concept of pipelining, introduction to Pipelined data path and control - Handling Data hazards & Control hazards.

Bachelor of Science in Data Science

(B. Sc. Data Science – First Year)

Programming Methodology And Data Structure

Unit I

Introduction to Programming - Program Concept, Characteristics of Programming, Stages in Program Development, Algorithms, Notations, Design, Flowcharts, Types of Programming Methodologies. Basics of C++: A Brief History of C++, Application of C++, Compiling & Linking, Tokens, Keywords, Identifiers & Constants Basic Data Types, User- Defined Data Types, Symbolic Constant, Type Compatibility Reference Variables, Operator in C++, Scope Resolution Operator, Member Dereferencing Operators, Memory Management Operators, Manipulators, Type Cast Operator. Functions In C++: The Main Function, Function Prototyping, Call by Reference Call by Address, Call by Value, Return by Reference, Inline Function, Default Arguments, Constant Arguments, Function Overloading, Function with Array.

Unit II

Classes & Objects: A Sample C++ Program with class, Defining Member Functions, Making an Outside Function Inline, Nesting of Member Functions, Private Member Functions, Arrays within a Class, Memory Allocation for Objects, Static Data Members, Static Member, Functions, Array of Objects, Object as Function Arguments, Friend Functions, Virtual functions, Returning Objects, Constant member functions, Pointer to Members, Local Classes.

Constructor & Destructor: Constructor, Parameterized Constructor, Multiple Constructors in a Class, Constructors with Default Arguments, Dynamic Initialization of Objects, Copy Constructor, Dynamic Constructor and Destructor.

Unit III

Inheritance: Defining Derived Classes, Single Inheritance, Making a Private Member Inheritable, Multilevel Inheritance, Hierarchical Inheritance, Multiple Inheritance, Hybrid Inheritance, Virtual Base Classes, Abstract Classes, Constructor in Derived Classes, Nesting of Classes. Operator Overloading & Type Conversion, Polymorphism, Pointers, Pointers with Arrays C++, Streams, C++ Stream Classes, Unformatted I/O Operation, Formatted I/O Operation, Managing Output with Manipulators, Exception Handling.

Unit IV

Data Structure: Basic concepts, Linear and Non-Linear data structures Algorithm Specification:
Introduction, Recursive algorithms, Data Abstraction, Performance analysis.Arrays: Representation of single, two-dimensional arrays, triangular arrays, sparse matrices-array and linked representations. Stacks: Operations, Array and Linked Implementations, Applications Infix to Postfix Conversion, Infix to Prefix Conversion, Postfix Expression Evaluation, Recursion Implementation.

Queues: Definition, Operations, Array and Linked Implementations. Circular Queue- Insertion and Deletion Operations, Dequeue (Double Ended Queue), Priority Queue- Implementation.

Unit V

Linked Lists: Singly Linked Lists, Operations, Concatenating, circularly linked lists- Operations for Circularly linked lists, Doubly Linked Lists- Operations, Doubly Circular Linked List, Header Linked List Trees: Representation of Trees, Binary tree, Properties of Binary Trees, Binary Tree Representations- Array and Linked Representations, Binary Tree Traversals, Threaded Binary Trees. Heap: Definition, Insertion, Deletion.

Bachelor of Science in Data Science

(B. Sc. Data Science – First Year)

Operating System

Unit I

Introduction to Operating System: What is Operating System? History and Evolution of OS, Basic OS functions, Resource Abstraction, Types of Operating Systems- Batch Systems, Multiprogramming Systems, Multiprocessing Systems, Time Sharing Systems, Distributed OS, Real time systems. Operating System for Personal Computers, Workstations and Hand-held Devices. Applications of various operating systems in real world. Some prevalent operating systems Windows, UNIX/Linux, Android, MacOS, Blackberry OS, Symbian, Bada etc.

UnitII

Process Management: Process Concepts, Process states & Process Control Block.
Process Scheduling: Scheduling Criteria, Scheduling Algorithms (Preemptive & Non Preemptive) - FCFS, SJF, SRTN, RR, Priority, Multiple-Processor, Real-Time, Multilevel Queue and Multilevel Feedback Queue Scheduling. Deadlock - Definition, Deadlock Characterization, Necessary and Sufficient Conditions for Deadlock. Deadlock Handling Approaches: Prevention, Avoidance, Detection and Recovery

Unit III

Memory Management: Introduction, Address Binding, Logical versus Physical Address Space, Swapping, Contiguous & Non-Contiguous Allocation, Fragmentation (Internal & External), Compaction, Paging, Segmentation, Virtual Memory, Demand Paging, Performance of Demand Paging, Page Replacement Algorithms.

File Management: Concept of File System(File Attributes, Operations, Types), Functions of File System, Types of File System, Access Methods (Sequential, Direct & other methods), Directory Structure (Single-Level, Two-Level, Tree-Structured, Acyclic-Graph, General Graph), Allocation Methods (Contiguous, Linked, Indexed)

Unit IV

Disk Management: Structure, Disk Scheduling Algorithms (FCFS, SSTF, SCAN, C-SCAN, LOOK), Swap Space Management, Disk Reliability, Recovery. Security: Security Threats, Security policy mechanism, Protection, Trusted Systems, Authentication and Internal Access Authorization, Windows Security.

Unit V

LINUX: Introduction, History and features of Linux, advantages, hardware requirements for installation, Linux architecture, file system of Linux - boot block, super block, inode table, data blocks. Linux standard directories, Linux kernel, Partitioning the hard drive for Linux, installing the Linux system, system - startup and shut-down process, init and run levels. Process, Swap Partition, fdisk, checking disk free spaces. Difference between CLI OS & GUI OS, Windows v/s Linux, Importance of Linux Kernel, Files and Directories. Concept of Open Source Software

Bachelor of Science in Data Science

(B. Sc. Data Science – First Year)

Introduction to Data Science & Python Programming

Unit I

Introduction to Data Science - Evolution of Data Science - Data Science Roles - Stages in a Data Science Project - Applications of Data Science in various fields - Data Security Issues. Data Collection and Data Pre-Processing Data Collection Strategies - Data Pre-Processing Overview - Data Cleaning - Data Integration and Transformation - Data Reduction - Data Discretization.

Unit II

Exploratory Data Analytics Descriptive Statistics - Mean, Standard Deviation, Skewness and Kurtosis - Box Plots - Pivot Table - Heat Map - Correlation Statistics - ANOVA.
Model Development Simple and Multiple Regression - Model Evaluation using Visualization - Residual Plot- Distribution Plot - Polynomial Regression and Pipelines - Measures for In-sample Evaluation - Prediction and Decision Making.

Unit III

Model Evaluation Generalization Error - Out-of-Sample Evaluation Metrics - Cross Validation - Overfitting - Under Fitting and Model Selection - Prediction by using Ridge Regression - Testing Multiple Parameters by using Grid Search.

Unit IV

Introduction to python language: Basic syntax, Literal Constants, Numbers, Variable and Basic data types, String, Escape Sequences, Operators and Expressions, Evaluation Order, Indentation, Input, Output, Functions, Comments. Data Structure: List, Tuples, Dictionary, Data Frame and Sets, constructing, indexing, slicing and content manipulation.

Unit V

Control Flow: Conditional Statements - If, If-else, Nested If-else. Iterative Statement - For, While, Nested Loops. Control statements - Break, Continue, and Pass.
Python-Functions: Syntax for defining a function, Calling a Function, Function Arguments, Anonymous Functions.

Bachelor of Science in Data Science

(B. Sc. Data Science – First Year)

Web Designing

Unit I

Introduction to Internet- World Wide Web, Internet Addressing, Browser, URL, Web server, website, homepage, Domain Name. Basic concepts. Softwares for Web Designing - Notepad/Notepad++, Dreamweaver, Blue Griffon, Net beans, Sea Monkey, Word press, Sublime.

Introduction to HTML: HTML Tags and Attributes, HTML Basic Tags, Formatting Tags, HTML Color Coding, Div and Span Tags for Grouping. Lists: Unordered Lists, Ordered Lists, Definition list. Images: Image and Image Mapping

Hyperlink: URL - Uniform Resource Locator, URL Encoding. Table: <table>, <th>, <tr>, <td>, <caption>, <thead>, <tbody>, <tfoot>, <colgroup>, <col>. Attributes Using Iframe as the Target

Form: <input>, <textarea>, <button>, <select>, <label>

Headers: Title, Base, Link, Styles, Script

HTML Meta Tag, XHTML, HTML Deprecated Tags & Attributes

Unit II

CSS: Introduction, Features and benefits of CSS, CSS Syntax, External Style Sheet using <link>, Multiple Style Sheets, Value Lengths and Percentages. Selectors: ID Selectors, Class Selectors, Grouping Selectors, Universal Selector, Descendant/Child Selectors, Attribute Selectors, CSS - Pseudo Classes. Color Background Cursor: background-image, background-repeat, background-position, CSS Cursor Text Fonts: color, background-color, text-decoration, text-align, vertical-align, text-indent, text transform, white-space, letter-spacing, word-spacing, line-height, font-family, font-size, font-style, font variant, font-weight.

Unit III

Lists Tables: list-style-type, list-style-position, list-style-image, list-style, CSS Tables (border, width & height, text-align, vertical-align, padding, color) Box Model: Borders & Outline, Margin & Padding, Height and width, CSS Dimensions. Display Positioning: CSS Visibility, CSS Display, CSS Scrollbars, CSS Positioning (Static Positioning, Fixed Positioning, Relative Positioning, Absolute Positioning), CSS Layers with Z-Index.Floats: The float Property, The clear Property, The clearfix Hack.

Unit IV

The JavaScript: Nature of JavaScript, Script Writing Basics, Enhancing HTML Documents with JavaScript, The Building Blocks.

Introduction to JavaScript, JavaScript Engines, Values, Variables and Operators, Variable Mutation, Basic Operators, Operator Precedence, JavaScript Types, Types Definition, Types in JavaScript, Objects, Type Conversion and Coercion, Static vs Dynamic Type Checking.

JavaScript Conditionals: Introduction to Conditionals, Conditionals in JavaScript, Ternary Operators and Conditionals. Conditional Ladder & Switch statement.

JavaScript Arrays: Introduction to Arrays, Declaring and Mutating Arrays, Array Methods and Properties, Replication with Array Methods, Multi-dimensional Arrays.

Unit V

JavaScript Loops: Introduction to Loops, Loops in JavaScript, While and Do/While Loops, For Loops, Break and Continue in Loops, Iterating Arrays, Iterating Objects.

JavaScript Functions: Introduction to Functions, Functions in JavaScript, Nested Functions in JavaScript, Arrow Functions in JavaScript, Function as an Argument, Function as the Returned Object, JavaScript Scope: Scope Introduction, Scope in JavaScript, Lexical Scope, Module Scope.

Method of Adding Interactivity to a Web Page, Creating Dynamic Web Pages; Concept of Java Scripting the Forms. Java Scripting the Forms, Basic Script Construction, Talking to the Form Objects, Organizing the Objects and Scripts, Field-Level Validation, Check Required Fields like Validating Zip Code, Automated Formatting, Format Phone, Format Money, Automatic Calculation, Calculate Expiration Date, Calculate Amount etc.

Bachelor of Science in Data Science

(B. Sc. Data Science – Second Year)

DBMS and Data Mining

UNIT- I

Introduction: Advantage of DBMS approach, various view of data, data independence, schema and subschema, primary concepts of data models, Database languages, transaction management, Database administrator and users, data dictionary, overall system architecture. ER model: basic concepts, design issues, mapping constraint, keys, ER diagram, weak and strong entity sets, specialization and generalization, aggregation, inheritance, design of ER schema, reduction of ER schema to tables.

UNIT II

Domains, Relations and Keys: domains, relations, kind of relations, relational database, various types of keys, candidate, primary, alternate and foreign keys. Relational Algebra & SQL: Features of good relational database design, Codd's rule, The structure, relational algebra with extended operations, modifications of Database, idea of relational calculus, basic structure of SQL, set operations, aggregate functions, null values, nested sub queries, derived relations, views, modification of Database, join relations, DDL in SQL. PL/SQL programming: working with stored procedures, triggers, cursor Database Integrity: general idea. Integrity rules, domain rules, attribute rules, relation rules, Database rules, assertions, triggers, integrity and SQL.

UNIT III

Functional Dependencies and Normalization: basic definitions, trivial and non trivial dependencies, closure set of dependencies and of attributes, irreducible set of dependencies, introduction to normalization, non loss decomposition, FD diagram, first, second, third Normal forms, dependency preservation, BCNF, multivalued dependencies and fourth normal form, Join dependency and fifth normal form.

UNIT IV

Transaction, concurrency and Recovery: basic concepts, ACID properties, Transaction states, implementation of atomicity and durability, concurrent executions, basic idea of serializability, basic idea of concurrency control, basic idea of deadlock, failure classification, storage structure types, stable storage implementation, data access, recovery and atomicity

log based recovery, deferred Database modification, immediate Database modification, checkpoints. Distributed Database: basic idea, distributed data storage, data replication, data fragmentation: horizontal, vertical and mixed fragmentation.

UNIT V

Emerging Fields in DBMS: object oriented Databases-basic idea and the model, object structure, object class, inheritance, multiple inheritance, object identity, data warehousing terminology, definitions, characteristics, data mining and it's overview, Database on www, multimedia Databases-difference with conventional DBMS, issues, similarity based retrieval, continuous media data, multimedia data formats, video

servers. Storage structure and file

organizations: overview of physical storage media, magnetic disk performance and optimization, basic idea of RAID, file organization, organization of records in files, basic concepts of indexing, ordered indices, basic idea of B-tree and B+-tree organization Network and hierarchical models: basic idea, data structure diagrams, DBTG model, implementations, tree structure diagram, implementation techniques, comparison of the three models.

UNIT –V I

Motivation, importance, Data type for Data Mining : relation Databases, Data Warehouses, Transactional databases, advanced database system and its applications, Data mining Functionalities: Concept/Class description, Association Analysis classification & Prediction, Cluster Analysis, Outlier Analysis, Evolution Analysis, Classification of Data Mining Systems, Major Issues in Data Mining.

UNIT – VII

Data Warehouse and OLAP Technology for Data Mining: Differences between Operational Database Systems and Data Warehouses, a multidimensional Data Model, Data Warehouse Architecture, Data Warehouse Architecture, Data Warehouse Implementation, Data Cube Technology.

Recommended Books:

1. A Silberschatz, H.F Korth, Sudersan “Database System Concepts” –, MGH Publication. 2.

C.J Date “An introduction to Database Systems” –6th ed.

3. Elmasri & Navathe “Fundamentals of Database systems” – III ed.

4. B.C. Desai. “An introduction to Database systems” BPB

5. Raghurama Krishnan “Database Systems” TMH

6. J. Han and M. Kamber, “Data Mining: Concepts and Techniques”, Morgan Kaufmann Pub. 7. Berson “Dataware housing, Data Mining & DLAP, @004, TMH.

8. W.H. Inmon “ Building the Datawarehouse, 3ed, Wiley India.

9. Anahory, “Data Warehousing in Real World”, Pearson Education.

10. Adriaans, “Data Mining”, Pearson Education. 6. S.K. Pujari, “Data Mining Techniques”,

University Press, Hyderabad

Bachelor of Science in Data Science

(B. Sc. Data Science – Second Year)

Big Data and Big Data Analysis

UNIT I

INTRODUCTION TO DATA SCIENCE AND BIG DATA: Introduction to Data Science – Data Science Process – Exploratory Data analysis – Big data: Definition, Risks of Big Data, Structure of Big Data – Web Data: The Original Big Data – Evolution Of Analytic Scalability – Analytic Processes and Tools – Analysis versus Reporting – Core Analytics versus Advanced Analytics– Modern Data Analytic Tools – Statistical Concepts: Sampling Distributions – Re Sampling – Statistical Inference – Introduction to Data Visualization.

UNIT II

DATA ANALYSIS USING R Univariate Analysis: Frequency, Mean, Median, Mode, Variance, Standard Deviation, Skewness and Kurtosis – Bivariate Analysis: Correlation – Regression Modeling: Linear and Logistic Regression – Multivariate Analysis – Graphical representation of Univariate, Bivariate and Multivariate Analysis in R: Bar Plot, Histogram, Box Plot, Line Plot, Scatter Plot, Lattice Plot, Regression Line, Two-Way cross Tabulation.

UNIT III

DATA MODELING Bayesian Modeling – Support Vector and Kernel Methods – Neuro – Fuzzy Modeling – Principal Component Analysis – Introduction to NoSQL: CAP Theorem, MongoDB: RDBMS VsMongoDB, Mongo DB Database Model, Data Types and Sharding – Data Modeling in HBase: Defining Schema – CRUD Operations

UNIT IV

DATA ANALYTICAL FRAMEWORKS Introduction to Hadoop: Hadoop Overview – RDBMS versus Hadoop – HDFS (Hadoop Distributed File System): Components and Block Replication – Introduction to MapReduce – Running Algorithms Using MapReduce – Introduction to HBase: HBase Architecture, HLog and HFile, Data Replication – Introduction to Hive, Spark and Apache Sqoop.

UNIT V

STREAM ANALYTICS Introduction To Streams Concepts – Stream Data Model and Architecture – Stream Computing – Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window

UNIT VI

INTRODUCTION TO BIG DATA Evolution of Big data – Best Practices for Big data Analytics – Big data characteristics – Validating – The Promotion of the Value of Big Data – Big Data Use Cases- Characteristics of Big Data Applications – Perception and

Quantification of Value -Understanding Big Data Storage – A General Overview of High Performance Architecture – HDFS – MapReduce and YARN – Map Reduce Programming Model

UNIT VII

CLUSTERING AND CLASSIFICATION Advanced Analytical Theory and Methods: Overview of Clustering – K-means – Use Cases – Overview of the Method – Determining the Number of Clusters – Diagnostics – Reasons to Choose and Cautions .- Classification: Decision Trees – Overview of a Decision Tree – The General Algorithm – Decision Tree Algorithms – Evaluating a Decision Tree – Decision Trees in R – Naïve Bayes – Bayes' Theorem – Naïve Bayes Classifier.

Recommended Books:

1. Anand Rajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press,

2. David Loshin, "Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph", Morgan Kaufmann/El sevier Publishers, 2013.

3. EMC Education Services, "Data Science and Big Data Analytics: Discovering, Analyzing,Visualizing and Presenting Data", Wiley publishers,2015. 4. Bart Baesens, "Analytics in a Big Data World: The Essential Guide to Data Science and its Applications", Wiley Publishers.

5. DietmarJannach and Markus Zanker, "Recommender Systems: An Introduction", Cambridge University Press

6. Kim H. Pries and Robert Dunnigan, "Big Data Analytics: A Practical Guide for Managers "CRC Press,

Bachelor of Science in Data Science

(B. Sc. Data Science – Second Year)

Artificial Intelligence and Machine Learning

UNIT I

General Issues and Overview of AI: The AI problems, what is an AI technique, Characteristics of AI applications. Introduction to LISP programming: Syntax and numeric functions, Basic list manipulation functions, predicates and conditionals, input output and local variables, iteraction and recursion, property lists and arrays.

UNIT II

Problem Solving, Search and Control Strategies General problem solving, production systems, control strategies forward and backward chaining, exhaustive searches depth first breadth first search. Heuristic Search Techniques Hill climbing, branch and bound technique, best first search & A* algorithm, AND / OR graphs, problem reduction & AO* algorithm, constraint satisfaction problems.

UNIT III

Knowledge Representations First order predicate calculus, skolemization, resolution principle & unification, interface mechanisms, horn's clauses, semantic networks, frame systems and value inheritance, scripts, conceptual dependency.

UNIT IV

Natural Language processing Parsing techniques, context free grammer, recursive transitions nets (RNT), augmented transition nets (ATN), case and logic grammers, symantic analysis. Game playing Minimax search procedure, alpha-beta cutoffs, additional refinments. Planning Overview an example domain the block word, component of planning systems, goal stack planning, non linear planning.

UNIT V

Probabilistic Reasoning and Uncertainty Probability theory, bayes theorem and bayesian networks, certainty factor. Expert Systems Introduction to expert system and application of expert systems, various expert system shells, vidwan frame work, knowledge acquisition, case studies, MYCIN. Learning Rote learning, learning by induction, explanation based learning

UNIT VI

Introduction to machine learning, scope and limitations, regression, probability, statistics and linear algebra for machine learning, convex optimization, data visualization, hypothesis function and testing, data distributions, data pre-processing, data augmentation, normalizing data sets, machine learning models, supervised and unsupervised learning.

UNIT VII

Linearity vs non linearity, activation functions like sigmoid, ReLU, etc., weights and bias, loss function, gradient descent, multilayer network, back propagation, weight initialization, training, testing, unstable gradient problem, auto encoders, batch normalization, dropout, L1 and L2 regularization, momentum, tuning hyper parameters.

Recommended Books:

1. Elaine Rich and Kevin Knight “Artifical Intelligence” - Tata McGraw Hill. 2. “Artifical Intelligence” 4 ed. Pearson

.3. Dan W. Patterson “Introduction to Artifical Intelligence and Expert Systems”, Prentice India.

4. Nils J. Nilson “Principles of Artifical Intelligence”, Narosa Publishing House.

5. Clocksin & C.S.Melish “Programming in PROLOG”, Narosa Publishing House.

6. M.Sasikumar,S.Ramani etc. “Rule based Expert System”, Narosa Publishing House

Bachelor of Science in Data Science

(B. Sc. Data Science – Second Year)

Advance web technology

UNIT I

Responsive web design and introduction to Bootstrap : Bootstrap grid, bootstrap components and plugins.

UNIT II

XML Introduction to XML, Comparing XML with HTML, Describing the Structure of XML - Declaration, Elements, Attributes, Comments, CDATA, XML Entity References, Parsers ,Describing Document Type Definitions, Using XSLT with XML :xsl:template Element, xsl:apply-templates Element,xsl:import , xsl:include Element, Element,xsl:element Element, xsl:attribute Element, xsl:value-of Element, using Conditional Statements, Sorting Elements, XSLTfunctions, Creating Well-formed and Valid Documents.

UNIT III

Introduction to Ajax – AJAX Web Application Model, Working of AJAX, Asynchronous Data Transfer with XMLHttpRequest - Creating the XMLHttpRequest Object, XMLHttpRequest Properties, XMLHttpRequest Methods, Using the XMLHttpRequest Object in Different Browsers, Reading a File Synchronously, Reading a File Asynchronously, Performing Tasks Using the XMLHttpRequest Object, Integrating PHP and AJAX-Sending Data from a Web Application to a Server, Validating a Field Using AJAX and PHP

UNIT IV

Handling XML Data using PHP and AJAX-JavaScript, properties for Extracting with nodeValue, Accessing XML, Elements by Name, Accessing Attribute Values in XML Elements. Validating XML Documents in Ajax Applications Retrieving Data from a Database Using PHP and AJAX Consuming Web Services Using AJAX-Exploring Web Service Protocols-SOAP,Web Service Description Language, UDDI, REST, Consuming Web Services Using AJAX

UNIT V

jQuery-JavaScript DOM objects their methods and properties-Window, History, Location Document, Form etc. Fundamentals of jQuery, Loading and using jQuery, using jQuery Library files, Call-back functions, jQuery Selectors , jQuery Methods to Access HTML Attributes, jQuery Methods of traversing, jQuery Manipulators, jQuery Events, jQuery Effects, jQuery with AJAX.

Recommended Books:

1. Bootstrap: Responsive Web Development

2. XML: A Beginner's Guide by Steven Holzner

3. AJAX For Beginners , Ivan Bayross and Sharanam Shah, SPD 4. Web Development with jQuery (WROX) by Richard York 5. Learning PHP, MySQL & JavaScript with j Query, 6 CSS & HTML – by Robin Nixon ,SPD

7. Ajax in Action Dave Crane, Eric Pascarello, Darren James 8. Ajax for Dummies Steve Holzner,PhD, Wiley Publishing Inc. \

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

INTRODUCTION TO DATA SCIENCE AND R PROGRAMMING

UNIT I:

Defining Data Science and Big data, Benefits and Uses, facets of Data, Data Science Process. History and Overview of R, Getting Started with R, R Nuts and Bolts

UNIT II:

The Data Science Process: Overview of the Data Science Process-Setting the research goal, Retrieving Data, Data Preparation, Exploration, Modeling, data Presentation and Automation. Getting Data in and out of R, Using reader package, Interfaces to the outside world.

UNIT III:

Machine Learning: Understanding why data scientists use machine learning-What is machine learning and why we should care about, Applications of machine learning in data science, Where it is used in data science, The modeling process, Types of Machine Learning-Supervised and Unsupervised.

UNIT IV:

Handling large Data on a Single Computer: The problems we face when handling large data, General Techniques for handling large volumes of data, Generating programming tips for dealing with large datasets. Case study- Predicting malicious URLs(This can be implemented in R).

UNIT V:

Sub setting R objects, Vectorised Operations, Managing Data Frames with the dplyr, Control structures, functions, Scoping rules of R, Coding Standards in R, Loop Functions, Debugging, Simulation

TEXT BOOKS:

1. DavyCielen, Arno.D.B.Maysman, Mohamed Ali, “Introducing Data Science” ManningPublications, 2016.
2. Roger D. Peng, “R Programming for DataScience” Lean Publishing, 2015.

REFERENCE BOOKS:

1. Nina Zumel, John Mount, “Practical Data Science with R”, Manning Publications, 2014. 2. Tony Ojeda, Sean Patrick Murphy, Benjamin Bengfort, AbhijitDasgupta, “Practical Data Science Cookbook”, Packt Publishing Ltd., 2014.

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

DATA MINING CONCEPTS AND TECHNIQUES

UNIT I:

An idea on Data Warehouse, Data mining-KDD versus data mining, Stages of the Data Mining Process-Task primitives., Data Mining Techniques – Data mining knowledge representation.

UNIT II

Data mining query languages- Integration of Data Mining System with a Data Warehouse Issues, Data pre-processing – Data Cleaning, Data transformation – Feature selection – Dimensionality reduction

UNIT III

Concept Description: Characterization and comparison What is Concept Description,Data Generalization by Attribute-Oriented Induction(AOI), AOI for Data Characterization, Efficient Implementation of AOI.

Mining Frequent Patterns, Associations and Correlations: Basic Concepts, FrequentItemset Mining Methods: Apriori method, generating Association Rules, Improvingthe Efficiency of Apriori, Pattern-Growth Approach for mining Frequent Item sets.

UNIT-IV

Classification Basic Concepts: Basic Concepts, Decision Tree Induction: Decision TreeInduction Algorithm, Attribute Selection Measures, Tree Pruning. Bayes Classification Methods.

UNIT-V

Classification by Back Propagation: Multi_Layer Feed Forward Neural Network. Support Vector Machines: Cases when the data are linearly separable and linearlyinseparable. Cluster Analysis: Cluster Analysis, Partitioning Methods, Hierarchal methods, Density based methods-DBSCAN.

TEXT BOOKS:

1. Jiawei Han, MichelineKamber, Jian Pei."Data Mining: Concepts and Techniques", 3rd Edition,Morgan Kaufmann Publishers, 2011.

2. AdelchiAzzalini, Bruno Scapa, "Data Analysis and Data mining" , 2ndEdiiton, Oxford Univeristy Press Inc., 2012.

REFERENCES BOOKS:

1. Alex Berson and Stephen J. Smith, "Data Warehousing, Data Mining & OLAP", 10th Edition, TataMcGraw Hill Edition , 2007.

2. G.K. Gupta, “Introduction to Data Mining with Case Studies”, 1st Edition, Easter Economy Edition, PHI, 2006.

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

PYTHON PROGRAMMING FOR DATA ANALYSIS

UNIT I:

What is Data Analysis? Differences between Data Analysis and Analytics, What is Python, Why Python for Data Analysis? What is Library, Essential Python Libraries. Python Language basics, I Python and Jupyter Notebook. Python Language Basics.

UNIT II:

Built-in Data Structures, Functions, Files and Operating System. NumPy Basics: Arrays and Vectorized Computation, The Numpynd array, Universal Functions, Array-Oriented Programming with Arrays, File Input and Output with Arrays, Linear Algebra, Pseudorandom Number Generation.

UNIT III:

Getting Started with Pandas: Introduction to Pandas Data Structures, Essential Functionality, Summarizing and Computing Descriptive Statistics

Data Loading, Storage and File Formats: Reading and Writing Data in Text Format, Binary Data Formats, Interacting with Web APIs, Interacting with Databases.

UNIT IV:

Data Cleaning and Preparation: Handling Missing Data, Data Transformation, String Manipulation.

Data Wrangling: Join, Combine and Reshape: Hierarchical Indexing, Combining and Merging Datasets, Reshaping and Pivoting.

UNIT V:

Introduction to Modeling Libraries in Python: Interfacing between pandas and Model code, Creating model descriptions with Patsy, Introduction to stats models.

Plotting and Visualization: A brief matplotlib API Primer, Plotting with Pandas and Seaborn, Other Python visualization tools.

TEXT BOOKS:

1. Wes McKinney “Python for Data Analysis” O’reilly Publications Second edition 2.
Charles R Suverance “Python for Everybody” Exploring data using Python 3

REFERENCE BOOKS:

3. John Zelle Michael Smith Python Programming, second edition 2010

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

BIG DATA ANALYTICS USING SPARK

UNIT I:

Introduction to Big Data: What is Big Data-Characteristics, Data in the Warehouse and Data in Hadoop, Why is Big Data Important- When to consider Big Data Solution, Applications. Introduction to Hadoop: Hadoop- definition, Application development in Hadoop. The building blocks of Hadoop, Name Node, Data Node, Secondary Name Node, Job Tracker and Task Tracker.

UNIT II:

Introduction to Spark: What is Apache Spark, Why Spark when Hadoop is there, Spark Features, , Spark components, Spark program flow, Spark Eco System. Differences between implementation of programs in Hadoop and Spark Programming environments.

UNIT III:

Spark Fundamentals- Using spark in action VM, Using Spark Shell and writing first spark program, Basic RDD actions and transformations.

Spark SQL-Working with Data Frames, Using SQL Commands, Saving and loading Data Frame.

UNIT IV:

Streaming in Spark- Writing spark streaming applications, Using external data sources, structured streaming.

Spark MLlib-Introduction to Machine Learning. Definition of Machine Learning, Machine Learning with Spark.

UNIT V:

Graph Representation in MapReduce: Graph Processing with Spark, Spark GraphX, GraphX features, Graph Examples, Graph algorithms-Shortest Path Algorithm.

TEXT BOOKS:

1. Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data by Dirk deRoos, Chris Eaton, George Lapis, Paul Zikopoulos, Tom Deutsch, 1st Edition, TMH,2012.

2. Spark in Action PetarZecevic, markoBonaci Manning Publications-2016. 3. Learning Spark“Holden KarauA. Konwinskietc.,”O’reilly Publications.

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

DATA VISUALIZATION

UNIT I:

Creating Visual Analytics with tableau desktop, connecting to your data-How to Connect to your data, What are generated Values? Knowing when to use a direct connection, Joining tables with tableau, blending different data sources in a single worksheet.

UNIT II:

Building your first Visualization- How Me works- Chart types, Text Tables, Maps, bar chart, Line charts, Area Fill charts and Pie charts, scatter plot, Bullet graph, Gantt charts, Sorting data in tableau, Enhancing Views with filters, sets groups and hierarchies.

UNIT III:

Creating calculations to enhance your data- What is aggregation, what are calculated values and table calculations, Using the calculation dialog box to create, Building formulas using table calculations, Using table calculation functions

UNIT IV:

Using maps to improve insights-Create a Standard Map View, Plotting your own locations on a map, Replace Tableau's standard maps, Shaping data to enable Point-to-Point mapping.

UNIT V:

Developing an Adhoc analysis environment- generating new data with forecasts, providing self evidence adhoc analysis with parameters, Editing views in tableau Server.

TEXT BOOKS:

1. Tableau your data-Daniel G. Murray and the Inter works BI team, Wiley Publications 2. Tableau Data Visualizaton Cookbook, AshutoshNandeshwar, PACKT publishing. 3. Storytelling with Data: A Data Visualization Guide for Business Professionals by Cole NussbaumerKnaflic (2014)

4. ggplot2: Elegant Graphics for Data Analysis by Hadley Wickham (2009)

REFERENCE BOOKS:

5. Designing Data Visualizations: Representing Informational Relationships by Noah Iliinsky, Julie Steele (2011)

6. Alexandru C. Telea – "Data Visualization principles and practice" Second Edition, CRC Publications

7. Joshua N. Millign–" Learning Tableau -2019" – Third Edition- Packt publications

Bachelor of Science in Data Science

(B. Sc. Data Science – Third Year)

Data Analytics with Tableau

UNIT I:

Introduction to Data Analytics: Big Data and Data Science, Big Data Architectures, A Short Taxonomy of Data Analytics, Examples of Data Use, History on Methodologies for Data Analytics. Descriptive Statistics: Scale Types, Descriptive Univariate Analysis, Descriptive Bivariate Analysis.

UNIT II:

Descriptive Multivariate Analysis: Multivariate Frequencies, Multivariate Data Visualization, Multivariate Statistics, Infographics and Word Clouds Data Quality and Preprocessing: Data Quality, converting to a Different Scale Type, Converting to a Different Scale, Data Transformation, Dimensionality Reduction.

UNIT III:

Clustering: Distance Measures, Clustering Validation, Clustering Techniques. Frequent Pattern Mining: Frequent Itemsets, Association Rules, Behind Support and Confidence, Other Types of Pattern.

UNIT IV:

Regression: Predictive Performance Estimation, Finding the Parameters of the Model, Technique and Model Selection.

Classification: Binary Classification, Predictive Performance Measures for Classification, Distance-based Learning Algorithms, Probabilistic Classification Algorithms.

UNIT V:

Additional Predictive Methods: Search-based Algorithms, Optimization-based Algorithms. Advanced Predictive Topics: Ensemble Learning, Algorithm Bias, Non-binary Classification Tasks, Advanced Data Preparation Techniques for Prediction.

TEXT BOOKS:

1. "A General Introduction to Data Analytics" by João Mendes Moreira, André C. P. L. F. de Carvalho, TomášHorváth, 2019 Edition, Wiley Publications.
2. "Data Analytics: Principles, Tools and Practices" by Dr. Gaurav Aroraa, ChitraLele, Dr. Munish Jindal, 2022 Edition, pbp publications
3. "Data Analytics" by Anil Maheshwari, First Edition, McGraw Hill Education

Bachelor of Science in Data Science

(B. Sc. Data Science – Fourth Year)

AI Concepts and Techniques with Python

UNIT I:

Problems and Search: What is Artificial Intelligence, The AI Problems, and Underlying Assumption, what is an AI Technique.

Problems, Problems Spaces, and Search: Defining the problem as a state space search, production systems, problems characteristics, issues in the design of search programs.

UNIT II:

Heuristic Search Techniques: Generate-and-test, Hill Climbing, Best-First Search, Problem Reduction, Constraint Satisfaction, Means-Ends Analysis

UNIT III:

Knowledge Representation Issues: Representations and Mapping, Approaches to Knowledge Representation, The frame problem. Using Predicate Logic: Representing simple facts in logic, Representing Isa relationships, predicates, Resolution

UNIT IV:

Representing Knowledge using Rules: Procedural Vs Declarative knowledge, Logic Programming, Forward Vs Backward Reasoning, Matching, Control Knowledge

UNIT V:

Symbolic Reasoning under Uncertainty: Introduction to Non-monotonic Reasoning, Logics for Non-monotonic Reasoning, Implementation issues, Augmenting a Problem solver, implementation: DFS, BFS.

Statistical Reasoning: Probability and Bayes Theorem, Certainty Factors and Rule-Based Systems, Bayesian Networks, Dempster-Shafer Theory.

Bachelor of Science in Data Science

(B. Sc. Data Science – Fourth Year)

Supervised ML with Python

UNIT I:

Machine Learning Basics: What is machine learning? Key terminology, Key tasks of machine learning, How to choose right algorithm, steps in developing a machine learning, why python? Getting started with Numpy library

Classifying with k-Nearest Neighbors: The k-Nearest Neighbors classification algorithm, Parsing and importing data from a text file, Creating scatter plots with Matplotlib, Normalizing numeric values

UNIT II:

Splitting datasets one feature at a time-Decision trees: Introducing decision trees, measuring consistency in a dataset, using recursion to construct a decision tree, plotting trees in Matplotlib

UNIT III:

Classifying with probability theory-Naïve Bayes: Using probability distributions for classification, learning the naïve Bayes classifier, Parsing data from RSS feeds, using naïve Bayes to reveal regional attitudes

UNIT IV:

Logistic regression: Classification with logistic regression and the sigmoid function, Using optimization to find the best regression coefficients, the gradient descent optimization algorithm, Dealing with missing values in the our data

UNIT V:

Support vector machines: Introducing support vector machines, using the SMO algorithm for optimization, using kernels to “transform” data, Comparing support vector machines with other classifiers

TEXT BOOK:

1. Machine learning in action, Peter Harrington by Manning publications

Bachelor of Science in Data Science

(B. Sc. Data Science – Fourth Year)

Unsupervised ML with Python

UNIT I:

Unsupervised Learning: Clustering: k-means clustering algorithm, Improving cluster performance with post processing, Bisecting k-means, Example: clustering points on a map

UNIT II:

Association analysis : Apriori algorithm: Association analysis, The Apriori principle, Finding frequent item sets with the Apriori algorithm, Mining association rules from frequent item sets, uncovering patterns in congressional voting

UNIT III:

Finding frequent item sets: FP-growth –FP trees, Build FP-tree, mining frequent from an FP tree, finding co-occurring words in a Twitter feed, mining a click stream from a news site.

UNIT IV:

Principal component analysis: Dimensionality reduction techniques, using PCA to reduce the dimensionality of semiconductor manufacturing data

UNIT V:

Singular value decomposition: Applications of the SVD, Matrix factorization, SVD in Python, Collaborative filtering–based recommendation engines, a restaurant dish recommendation engine

TEXT BOOK:

1. Machine learning in action, Peter Harrington by Manning publications

Bachelor of Science in Data Science

(B. Sc. Data Science – Fourth Year)

Deep Learning Neural Networks With Python

UNIT I:

Introduction to Deep Learning: Artificial intelligence, machine learning and deep learning, history of machine learning, Why deep learning? Why now?

The mathematical building blocks of neural networks: A first look at a neural network, Data representations for neural networks, The gears of neural networks: tensor operations, The engine of neural networks: gradient-based optimization.

UNIT II:

Getting started with neural networks: Anatomy of a neural network, Introduction to Keras, Setting up a deep-learning workstation, Classifying movie reviews: a binary classification Example, Classifying newswires: a multiclass classification example, Predicting house prices: a regression example.

Fundamentals of machine learning: Four branches of machine learning, Evaluating machine learning models, Data preprocessing, feature engineering and feature learning, Overfitting and underfitting, The universal workflow of machine learning.

UNIT III:

Deep learning for computer vision: Introduction to convnets, Training a convnet from scratch on a small dataset, Using a pretrained convent, Visualizing what convnets learn.

UNIT IV:

Deep learning for text and sequences: Working with text data, Understanding recurrent neural networks, Advanced use of recurrent neural networks, Sequence processing with convnets.

UNIT V:

Advanced deep-learning best practices: Going beyond the Sequential model: theKeras functional API, Inspecting and monitoring deep-learning models using Keras callbacks and Tensor Board, Getting the most out of your models.

TEXT BOOKS:

1. "Deep Learning with Python" by Francois Chollet, , 2018 Edition, Manning Publications.

2. "Deep Learning with Python" by Nikhil Ketkar, JojoMoolayil, Second Edition, Apress. 3. "Python Deep Learning" by Ivan Vasilev, Daniel Slatter, Second Edition, Packt Publications.