# Sanjay Ghodawat University

## Kolhapur

## School of Engineering & Technology

## Department of Artificial Intelligence & Machine Learning

### M. Tech. Artificial Intelligence & Data Science

# Syllabus

(Programme Structure and Course Contents)

Academic Year 2024-25

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science
# (AY: 2024-25)

## Programme Structure for M. Tech. Artificial Intelligence & Data Science First Year

### Semester I: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSPC101 | Linear Algebra and Calculus | 3 | - | - | 3 | - | - | 3 |
| PDSPC102 | Cloud Computing | 3 | - | - | 3 | - | - | 3 |
| PDSPC103 | Scalable Systems for Data Science | 3 | - | - | 3 | - | - | 3 |
| PDSPE1XX | Program Specific Elective I | 3 | - | - | 3 | - | - | 3 |
| PSTRM101 | Research Methodology | 2 | - | - | - | 2 | - | 2 |
| PXXOE1XX | Open/Multidisciplinary Elective I | 2 | - | - | 2 | - | - | 2 |
| PDSVS101 | Natural language Processing | - | 4 | - | - | 2 | - | 2 |
| PDSPC104 | Open Source tools for Machine Learning & Data Visualization Lab | - | 4 | - | - | 2 | - | 2 |
| | **Total** | **16** | **08** | **-** | **14** | **6** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech. Artificial Intelligence & Data Science First Year
### Semester I: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSPC101 | Linear Algebra and Calculus | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPC102 | Cloud Computing | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPC103 | Scalable Systems for Data Science | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPE1XX | Program Specific Elective I | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PSTRM101 | Research Methodology | - | 10 | 5 | 15 | 35 | - | - | 50 |
| PXXOE1XX | Open/Multidisciplinary Elective I | - | 10 | 5 | 15 | 35 | - | - | 50 |
| PDSVS101 | Natural language Processing | - | - | - | - | - | 50 | - | 25 |
| PDSPC104 | Open Source tools for Machine Learning & Data Visualization Lab | - | - | - | - | - | 25 | 25 | 50 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, SEE - Semester End Examination**

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science
# (AY: 2024-25)

## Programme Structure for M. Tech. Artificial Intelligence & Data Science First Year

## Semester II: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSPC201 | Statistics for Signal Analysis | 3 | - | - | 3 | - | - | 3 |
| PDSPC202 | Optimization Techniques for Analytics | 3 | - | - | 3 | - | - | 3 |
| PDSPC203 | Deep Learning | 3 | - | - | 3 | - | - | 3 |
| PDSPE2XX | Program Specific Elective II | 3 | - | - | 3 | - | - | 3 |
| PSTRM201 | Research Publications and Ethics | 2 | - | - | 2 | - | - | 2 |
| PXXOE2XX | Open/Multidisciplinary Elective II | 2 | - | - | 2 | - | - | 2 |
| PDSVS201 | Web & Social Media Analytics | - | 4 | - | - | 2 | - | 2 |
| PDSPC204 | Advanced Data Visualization Lab | - | 4 | - | - | 2 | - | 2 |
| | **Total** | **16** | **08** | **-** | **14** | **6** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech. Artificial Intelligence & Data Science First Year
## Semester II: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSPC201 | Statistics for Signal Analysis | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPC202 | Optimization Techniques for Analytics | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPC203 | Deep Learning | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PDSPE2XX | Program Specific Elective II | 10 | 10 | 5 | 25 | 50 | - | - | 75 |
| PSTRM201 | Research Publications and Ethics | - | 10 | 5 | 15 | 35 | - | - | 50 |
| PXXOE2XX | Open/Multidisciplinary Elective II | - | 10 | 5 | 15 | 35 | - | - | 50 |
| PDSVS201 | Web & Social Media Analytics | - | - | - | - | - | 50 | - | 50 |
| PDSPC204 | Advanced Data Visualization Lab | - | - | - | - | - | 25 | 25 | 50 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, Term, SEE - Semester End Examination**

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science (AY: 2024-25)

**Exit Option:** Students who exit at the end of first year shall be awarded a PG Diploma after successful completion of first year with additional 4 credits earned through industry internship/field work/ vocational training/generic course/(s).

## Programme Structure for M. Tech Artificial Intelligence & Data Science Second Year Semester III: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSPC301 | MOOC I | 3 | - | - | 3 | - | - | 3 |
| PDSPC302 | MOOC II | 3 | - | - | 3 | - | - | 3 |
| PDSPR301 | Dissertation Phase I | - | - | - | - | 6 | - | 6 |
| PDSOT301 | Industry Internship | - | - | - | - | 8 | - | 8 |
| | **Total** | **6** | **-** | **-** | **06** | **14** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech Artificial Intelligence & Data Science Second Year Semester III: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSPC301 | MOOC I | - | - | - | - | 50 | 25 | - | 75 |
| PDSPC302 | MOOC II | - | - | - | - | 50 | 25 | - | 75 |
| PDSPR301 | Dissertation Phase I | - | - | - | - | - | 100 | 50 | 150 |
| PDSOT301 | Industry Internship | - | - | - | - | - | 150 | 50 | 200 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, SEE - Semester End Examination**

# OR

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science
# (AY: 2024-25)

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year

## Semester III: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSRP301 | Dissertation Phase I – Research | - | - | - | - | 20 | - | 20 |
| | **Total** | **-** | **-** | **-** | **-** | **20** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year
## Semester III: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSRP301 | Dissertation Phase I – Research | - | - | - | - | - | 350 | 150 | 500 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, SEE - Semester End Examination**

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science
# (AY: 2024-25)

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year
## Semester IV: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSPR401 | Dissertation Phase II | | - | - | - | 20 | - | 20 |
| | **Total** | **-** | **-** | **-** | **-** | **20** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year
## Semester IV: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSPR401 | Dissertation Phase II | | | | | - | 350 | 150 | 500 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, SEE - Semester End Examination**

# OR

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science
# (AY: 2024-25)

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year
## Semester IV: Teaching Scheme

| Course Code | Course Name | Teaching Scheme (Hrs/week) | | | Credits Assigned | | | |
|---|---|---|---|---|---|---|---|---|
| | | Th | Pr | Tu | Th | Pr | Tu | Total credits |
| PDSRP401 | Dissertation Phase II – Research | | - | - | - | 20 | - | 20 |
| | **Total** | **-** | **-** | **-** | **-** | **20** | **-** | **20** |

**Th: Theory, Pr: Practical, Tu: Tutorial**

## Programme Structure for M. Tech. Artificial Intelligence & Data Science Second Year
## Semester IV: Evaluation Scheme

| Course Code | Course Name | *Theory Marks* | | | | | *Practical Marks* | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | | Internal Assessment | | | | SEE | Term work | Practical Oral/POE | |
| | | T1 | T2 | FET | Total | | | | |
| PDSRP401 | Dissertation Phase II - Research | | | | | - | 350 | 150 | 500 |

**FET – Faculty evaluation for Theory, T1, T2, Continuous Assessment Test, SEE - Semester End Examination**

# Curriculum Structure of M. Tech. Artificial Intelligence & Data Science (AY: 2024-25)

## Programme structure for M. Tech. Artificial Intelligence & Data Science

### Open/Multidisciplinary Elective

| Sr.No | Course Code | Open/Multidisciplinary Elective- I | Credits |
|---|---|---|---|
| 1. | PDSOE1XX | Business Intelligence | 2 |
| **Open/Multidisciplinary Elective- II** | | | |
| 2. | PDSOE2XX | Big Data Analytics | 2 |

## Programme structure for M. Tech. Artificial Intelligence & Data Science

### Program Specific Elective

| Sr.No | Course Code | Program Specific Elective - I | Credits |
|---|---|---|---|
| 1. | PDSPE101 | Generative AI | 3 |
| 2. | PDSPE102 | Computational Intelligence | 3 |
| **Program Specific Elective - II** | | | |
| 1. | PDSPE201 | Bioinformatics | 3 |
| 2. | PDSPE202 | Pattern Recognition | 3 |

# Proposed Syllabus for

# M. Tech. Artificial Intelligence & Data Science

# As per NEP2020 Guidelines

# AY 2024-25

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPC101 | Linear Algebra and Calculus | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPC101 | Linear Algebra and Calculus | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** Linear algebra and Basic Analysis is the study of linear systems of equations, vector spaces, and linear transformations. Solving systems of linear equations is a basic tool of many mathematical procedures used for solving problems in science and engineering.

**Pre-requisites**: Fundamental knowledge of linear algebra and calculus is required.

**Course Objectives**:

1. To Use computational techniques and algebraic skills essential for the study of linear systems equations, matrix algebra, vector spaces
2. To Compute Eigen values and Eigen vectors find orthogonality and diagonalization

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Develop** multivariable functions that model real-world phenomena and solve related calculus problems.

**CO2** **Examine** the effect of linear transformations on vector spaces.

**CO3** **Analyze** the intermediate steps and results of the Gram-Schmidt algorithm & matrices.

**CO4** **Evaluate** Advanced Matrix Concepts.

**CO5** **Determine** applications or algorithms that utilize Eigen Theory

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Calculus** | |
| 1 | 1.1 | Revision of basic Calculus (Integration, differentiation). Introduction to Multivariate Calculus. Introduction to Taylor Series, Solution of linear equations, Vector space | 9 |
| | | | |
| 2.0 | | **Advanced Linear Algebra** | |
| 2 | 2.1 | Linear transformation, matrix representation, inner-products and norms Orthogonality | 9 |
| | | | |
| 3.0 | | **Matrix Theory** | |
| 3 | 3.1 | Gram-Schmidt algorithm , Matrices and Basic Operations, interpretation of matrices as linear mappings | 9 |
| | | | |
| 4.0 | | **Advanced Matrix Concepts** | |
| 4 | 4.1 | Properties of determinants, singular and non-singular matrices, examples, Matrix Inverse, finding an inverse matrix, Characteristic Polynomial | 9 |
| | | | |
| 5.0 | | **Eigen Theory** | |
| 5 | 5.1 | Eigen values and eigenvectors, Definition of Left/right Eigen values and Eigenvectors | 9 |
| | 5.2 | Caley-Hamilton theorem, interpretation of eigenvalues/vectors, Singular Value Decomposition | |

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Text\Reference Books**

1 Strang, Linear Algebra and its Applications, Wellesley-Cambridge Press, 2016.
2 S. Axler, Linear Algebra Done Right, Springer, 2015

- Internal Assessment (T1, T2 and FET):

  1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.
  2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

  1. Question paper will comprise of 5 questions, each carrying 10 marks.
  2. The duration of end semester examination shall be 2 hours.
  3. The students need to solve all 5 questions.
  4. Question No.1 will be compulsory and based on entire syllabus.
  5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPC102 | Cloud Computing | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPC102 | Cloud Computing | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** This course provides an in-depth exploration of advanced topics in cloud computing, including cloud architecture, service models, deployment strategies, security, and cutting-edge trends. Students will gain hands-on experience with leading cloud platforms and develop skills to design, implement, and manage robust cloud-based solutions.

**Pre-requisites**: Basic knowledge of operating system, computer networking.

**Course Objectives:**

1.To understand advanced cloud computing concepts and architectures.
2. To explore various cloud service models and deployment strategies.
3.To gain expertise in cloud security, compliance, and governance.
4.To learn about advanced cloud services, including containerization, orchestration, and server less computing.
5.To apply cloud computing knowledge to real-world projects and case studies.

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Determine** core principles, architecture, and models of cloud computing.

**CO2** **Deploy** cloud-based applications using leading cloud service providers

**CO3** **Identify** and implement cloud security best practices, including identity and access management, data protection, and encryption

**CO4** **Assess** the security and compliance aspects of cloud storage solutions in relation to organizational requirements.

**CO5** **Explore** emerging trends such as edge computing, IoT, AI/ML integration, and blockchain in the cloud.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Introduction and Cloud Architecture:** | |
| 1 | 1.1 | Overview of cloud computing concepts and models, Cloud computing architecture: NIST reference architecture, Cloud service models: IaaS, PaaS, SaaS, Cloud deployment models: Public, Private, Hybrid, Community | 9 |
| | | | |
| 2.0 | | **Advanced Cloud Services** | |
| 2 | 2.1 | Containerization and Docker, Container orchestration with Kubernetes, Server less computing and Function as a Service (FaaS), Micro services architecture and its implementation in the cloud | 9 |
| | | | |
| 3.0 | | **Cloud Security and Compliance** | |
| 3 | 3.1 | Cloud security fundamentals and best practices, Identity and access management (IAM), Data protection, encryption, and key management ,Compliance and regulatory issues in cloud computing (GDPR, HIPAA, etc.),Incident response and disaster recovery planning | 9 |
| | | | |
| 4.0 | | **Cloud Storage and Data Management** | |
| 4 | 4.1 | Cloud storage architectures and services, Data management in the cloud: Databases, data lakes, and data warehouses, Big Data analytics and machine learning in the cloud,Case studies on cloud-based data solutions | 9 |
| | | | |
| 5.0 | | **Emerging Trends in Cloud Computing** | |
| 5 | 5.1 | Edge computing and its integration with the cloud, Internet of Things (IoT) and cloud convergence, Artificial Intelligence (AI) and Machine Learning (ML) in the cloud,Block chain and cloud computing | 9 |

**Text Books**

1.Cloud Computing: Concepts, Technology & Architecture" by Thomas Erl, Zaigham Mahmood, and Ricardo Puttini

2. "Architecting the Cloud: Design Decisions for Cloud Computing Service Models (SaaS, PaaS, and IaaS)" by Michael J. Kavis

**Reference:**

1."Cloud Computing: Theory and Practice" by Dan C. Marinescu

2. "Mastering Cloud Computing: Foundations and Applications Programming" by Rajkumar Buyya, Christian Vecchiola, and S. Thamarai Selvi

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| PDSPC103 | Scalable Systems for Data Science | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| | | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| PDSPC103 | Scalable Systems for Data Science | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| | | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** This course will teach the fundamental "systems" aspects of designing and using scalable data science platforms. Such platforms are used to acquire, store and query large, fast and linked datasets, to train machine learning models, and to process and analyze large datasets.

**Pre-requisites**: Data Structures, Programming and Algorithm concepts. Programming.

**Course Objectives:**

1. To introduce systems and approaches for large scale data science problems
2. To understand handling large datasets
3. To learn how large scale machine learning and distributed machine learning approaches work

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Explain** different data mining techniques & hash functions.

**CO2** **Summarize** various techniques for set operations and representations

**CO3** **Implement** different Data Mining Techniques

**CO4** **Design** a clustering solution for a specific data set

**CO5** **Develop** a large-scale machine learning solution

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Introduction** | |
| 1 | 1.1 | Overview of data mining and map-reduce, Hash Functions- Indexes, Shingling LSH, Mining Data Streams-Finding similarities near-neighbor search, shingling of documents, Similarity-Preserving | 9 |
| | | | |
| 2.0 | | **Data Structures and Web Analysis** | |
| 2 | 2.1 | Summaries of Sets, Locality-Sensitive Hashing for Documents, Distance Measures, Link-analysis Page Rank, Link spam, Hubs and authorities | 9 |
| | | | |
| 3.0 | | **Data Mining Techniques:** | |
| 3 | 3.1 | Frequent Item sets Market based model, A-Priori Algorithm, Handling larger data sets in memory, Limited-pass algorithms | 9 |
| | | | |
| 4.0 | | **Clustering** | |
| 4 | 4.1 | Clustering Hierarchical clustering, k-means, CURE, Clustering in Non-Euclidean Spaces, Clustering for Streams and Parallelism | 9 |
| | | | |
| 5.0 | | **Web Advertising and Recommendation Systems** | |
| 5 | 5.1 | Advertising on the web Matching problem, ad-words problem, Recommendation Systems-Content- Based Recommendations, Collaborative Filtering, Dimensionality Reduction | 9 |
| | 5.2 | Large-scale machine learning, Parallel Implementation of Perceptrons, Parallel implementation of SVM, Dealing with High-Dimensional Euclidean Data in nearest neighbors, Distributed machine learning | |

**Text Books**

1.Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman, Mining of massive datasets, Cam-bridge University Press,2014

2.Jimmy Lin and Chris Dyer, Data-Intensive Text Processing with MapReduce, First edition,Morgan and Claypool Publishers,2010

**Reference:**

1.Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills, Advanced Analytics with Spark: Patterns for Learning from Data at Scale,Oreilly, 2015

2.Ankit Jain, Mastering Apache Storm: Processing big data streaming in real time, Packt Publishing,2017.

- Internal Assessment (T1, T2 and FET):

1.T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2.Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPE101 | Program Specific Elective I- Generative AI | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPE101 | Program Specific Elective I- Generative AI | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:**

This course provides a deep dive into Generative AI, covering neural network-based language models, advanced Transformers, and diverse generative models including GANs and autoregressive models. It explores applications like ChatGPT and ethical considerations such as bias and privacy, preparing students to navigate the complexities of AI responsibly..

**Pre-requisites**: Mathematical Foundations and machine learning.

**Course Objectives:**

- Explain the core principles and concepts of generative artificial intelligence.
- Explore various types of generative modelling techniques.
- Develop skills in implementing and training generative models.
- Encourage critical thinking about the ethical implications of generative AI.

**Course Outcomes:** after the end of this course students will be able to

**CO1** Define and explain the core concepts, principles, and applications of generative AI.

**CO2** Analyze and compare different generative modeling techniques.

**CO3** Design basic generative models.

**CO4** Evaluate the performance of generative models.

**CO5** Identify the ethical implications of generative AI

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

## Course Contents

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Overview of Generative AI:** | |
| 1 | 1.1 | Introduction to Neural Network based Language Models- Recurrent Neural Networks, Gated Recurrent Unit, Encoder-Decoder Networks, Probability and Statistics for Generative AI- Bayesian networks, Markov chains. | 9 |
| | | | |
| 2.0 | | **Transformers and Large Language Models (LLMs):** | |
| 2 | 2.1 | Language Models, Transformer Architecture, Motivation for Transformer, Architecture, Encoder-Decoder Architecture, Attention, Position-wise Feed-Forward Networks, Advantages and Limitations of Transformer Architecture. | 9 |
| | | | |
| 3.0 | | **Networks and Models:** | |
| 3 | 3.1 | Auto encoders, Variation Auto encoders, latent space, Generative Adversarial Networks (GANs)- Deep Convolution GAN (DCGAN),Wasserstein GAN, Conditional GAN, Autoregressive Models- Long Short-Term Memory Network (LSTM), Diffusion Models- Types of Diffusion Models, Architecture, Latent Diffusion Model (LDM), Benefits and Significance. | 9 |
| | | | |
| 4.0 | | **Applications and Ethical Implications:** | |
| 4 | 4.1 | Applications - ChatGPT Architecture, Google Bard, Claude 2, Falcon AI, LLaMa 2, Dolly 2, DALL-E 2,Midjourney. Ethics- Bias and Fairness in Generative Models. | 9 |
| | | | |
| 5.0 | | **Ethical Considerations and Challenges in Generative AI** | |
| 5 | 5.1 | Bias and Fairness in Generative Models, Privacy and Security Concerns,Responsible AI Practices | 9 |

**Text Books**

1.Generative Deep Learning: A Practical Guide by David Foster

2. Applied Generative AI for Beginners: Practical Knowledge on Diffusion Models,ChatGPT, and Other LLMs by Akshay Kulkarni, Adarsha Shivananda, Anoosh Kulkarni and Dilip Gudivada

**Reference:**

1. Generative AI by Tom Taulli
2. Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville
3. Generative Adversarial Networks: An Introductory Guide by Luke Metz
4. Auto encoders: Neural Networks for Unsupervised Learning by Ian Goodfellow

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPE102 | Program Specific Elective I- Computational Intelligence | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPE102 | Program Specific Elective I- Computational Intelligence | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** Computational intelligence is a research field that studies to realize the intelligent human behaviors on a computer. The ultimate goal is to make a computer that can learn, plan, and solve problems autonomously. The research areas include problem solving, reasoning, planning, natural language understanding, computer vision, automatic programming, machine learning, and so on.

**Pre-requisites**: Foundation in calculus, linear algebra, and statistics

**Course Objectives:**

1. To provide the most fundamental knowledge to the students so that they can understand what the computational intelligence is.
2. To Explore use cases and applications of computational intelligence

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Explain** knowledge of the fundamental principles of artificial intelligent systems

**CO2** **Assess** the benefits and challenges of using problem reduction for different types of complex problems.

**CO3** **Design** a knowledge representation system

**CO4** **Implement** an expert system using a LISP-based shell & Prolog.

**CO5** **Aanlyze** Problem Decomposition & Experts Systems

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Introduction** | |
| 1 | 1.1 | AI problem; AI techniques, problem as a state space search, Production Systems, Issues in design of search programs. Physical Symbol Systems and the scope of Symbolic AI, Agents. State Space Search: Depth First Search, Breadth First Search. | 9 |
| | | | |
| 2.0 | | **Heuristic Search Techniques** | |
| 2 | 2.1 | Generate-And- Test, Hill Climbing, Best-First Search. Problem Reduction, Simulated Annealing, Constraint Satisfaction, Means-Ends Analysis. Population Based Search: Genetic Algorithms, Ant Colony Optimization. | 9 |
| | | | |
| 3.0 | | **Knowledge Representation** | |
| 3 | 3.1 | Procedural Vs Declarative Knowledge, Representations & Approaches to Knowledge Representation, Forward Vs Backward Reasoning, Matching Techniques, Partial Matching, RETE Matching Algorithms. | 9 |
| | | | |
| 4.0 | | **Logic Based Programming** | |
| 4 | 4.1 | AI Programming languages: Overview of LISP, Search Strategies in LISP, Pattern matching in LISP , An Expert system Shell in LISP, Over view of Prolog, Production System using Prolog | 9 |
| | | | |
| 5.0 | | **Problem Decomposition & Experts Systems** | |
| 5 | 5.1 | Algorithm AO, Game Playing, Game Playing:Algorithms Minimax, AlphaBeta, | 9 |
| | 5.2 | Overview of an Expert System, Structure of an Expert Systems, Different Types of Expert Systems- Rule Based, Model Based, Case Based and Hybrid Expert Systems. Knowledge Building SystemTools, Expert System Shells, Fuzzy Expert systems. | |

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Text Books**

1 George F Luger, *Artificial Intelligence*, Pearson Education Publications, 2001.
2 Elaine Rich and Knight, *Artificial Intelligence*, Mcgraw-Hill Publications, 2017

**Reference:**

1. David Rolston, *Principles of AI and Expert system development,* MGH, 1988.

- Internal Assessment (T1, T2 and FET):

1.T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2.Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PSTRM101 | Research Methodology | 02 | - | - | 02 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PSTRM101 | Research Methodology | | 10 | 5 | 15 | 40% | 35 | 40% | 50 |

**Course Description**

This course will provide an opportunity for participants to establish or advance their understanding of research through critical exploration of research language, ethics, and approaches. The course introduces the language of research, ethical principles and challenges, and the elements of the research process within quantitative, qualitative, and mixed methods approaches.

**Pre-requisites: NIL**

**Course Objectives**

- To understand the types of research and their applications and research process arising in many engineering problems.
- To Practice the ethical standards in research
- To Understand Computation of Jacobian of functions of several variables and their applications.

**Course Outcomes:** After the successful completion of the course students will able to:

- **CO1** Explain the types of research and their applications and research process
- **CO2** Demonstrate the ethical standards in research
- **CO3** Identify and state the research problem clearly
- **CO4** Design carry out research literature review and identify gaps in literature pertaining to a particular research issue.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | Research Fundamentals and Types | 9 |
| 1 | 1.1 | Research Fundamentals and Types:<br>Definition and objectives of Research: Research method and research methodology, Concept of Construct, Postulate, Proposition, Thesis, Hypothesis, Law, Principle. Philosophy and validity and reliability of research. Objective of research. Types of Research: Pure and applied research. Descriptive and explanatory research. Qualitative and quantitative approaches Ethical Issues in research | |
| 2.0 | | Research Process and Research Modelling | 9 |
| 2 | 2.1 | Steps in research process Formulating the Research Problem, Literature Review, Developing the objectives, Preparing the research design including sample Design Relevance, interest, available data, choice of data, Analysis of data, Generalization and interpretation, | |
| | | | |
| 3.0 | | Outcome of research | 9 |
| 3 | 3.1 | Outcome of Research: Design of sample surveys Sampling design, measurement and scaling, data collection ,data preparation and data analysis using various techniques Research outcomes Research papers and Publications | |
| | | | |
| 4.0 | | Hypothesis Formulation and Testing | 9 |
| 4 | 4.1 | Probability distributions. Type I and II error, testing of mean, proportion, tests for equality of mean and variances of two populations, confidence interval, Z test and $\chi 2$ test for goodness of fit, ANOVA (one-way classification), | |
| | | Non-parametric tests: sign test, U test. Correlation and Regression Analysis: Karl Pearson's and Rank Correlation coefficient, simple linear regression: least squares method | |
| | | | |
| 5.0 | | Interpretation and Report Writing: | 9 |
| 5 | 5.1 | Meaning Of Interpretation, Why Interpretation?, Technique Of Interpretation, Precautions In Interpretation, Significance Of Report Writing, Different Steps In Writing Report, Layout Of The Research Report, Mechanics Of Writing A Research Report, Precautions For Writing Research Report | |

Text Books

1. Kothari C R, Research Methodology (Methods and Techniques), 2004, New Age International Publishers, New Delhi.

**References**

1. Pannnerselvam, R Research Methodology, 2014, Phi Learning Private Limited, Delhi.
2. Saravanavel P, Research Methodology, 2011, Kitab Mahal, Allahabad.
3. Kumar R. Research Methodology, 2011, SAGE Publications Asia-Pacific Ptd Singapore.
4. Singh Yogesh Kumar, 2011, Fundamental of Research Methodology and Statitics, New Age International Publishers, New Delhi.
5. Michael V P 2000, Research Methodology in Management Himalaya Publishing House, Mumbai.

**Internal Assessment (T1, T2 and FET)**

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.
2. FET shall be assessed for 5 marks separately.

**End Semester Examination**

1. Question paper will be of 50 marks comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be Two hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| P70OE1XX | Open/Multidisciplinary Elective I- Business Intelligence | 02 | - | - | 02 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| P70OE1XX | Open/Multidisciplinary Elective I- Business Intelligence | | 10 | 5 | 15 | 40% | 35 | 40% | 50 |

**Course Description:**

Business Intelligence (BI) is a crucial component of modern enterprises, enabling data-driven decision-making and strategic insights. This course provides participants with a comprehensive understanding of Business Intelligence concepts, tools, and techniques. Through a combination of theoretical lectures, hands-on exercises, and real-world case studies, participants will learn how to leverage BI technologies to extract valuable insights from data and drive business growth.

**Pre-requisites:** Algorithms & Data Structures

**Course Objective:**

- To Understand the Fundamentals of Business Intelligence.
- To Explore the role of BI in facilitating data-driven decision-making and gaining strategic insights.
- To Explore data visualization best practices.
- To hands-on experience with popular BI tools like Microsoft Power BI

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Solve4** unstructured business problems using DSS

**CO2** **Demonstrate3** different mathematical models for decision making

**CO3** **Interpret3** data to understand relationships between the underlying business processes of an organization

**CO4** **Recommend2** different strategic decisions using multi-criteria decision making expert systems.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Contents | | | |
|---|---|---|---|
| **Module** | **Unit** | **Description** | **Hours** |
| 1.0 | | **Overview of business intelligence, analytics, and decision support** | |
| 1 | 1.1 | Information Systems Support for Decision Making, An Early Framework for Computerized Decision Support, The Concept of Decision Support Systems, A Framework for Business Intelligence, Business Analytics Overview, Brief Introduction to Big Data Analytics | 7 |
| | | | |
| 2.0 | | **Foundations and technologies for decision making** | |
| 2 | 2.1 | Decision Making: Introduction and Definitions, Phases of the Decision Making Process, The Intelligence Phase, Design Phase, Choice Phase, Implementation Phase, Decision Support Systems: Capabilities, Classification, Components. | 7 |
| | | | |
| 3.0 | | **Optimization and Multi-Criteria Decision Making Systems** | |
| 3 | 3.1 | Decision Support Systems Modelling, Structure of Mathematical Models for Decision Support, Decision Making Under Certainty, Uncertainty and Risk. Decision Modelling with Spreadsheets, Mathematical Programming Optimization Decision Analysis with Decision Tables and Decision Trees Multi-Criteria Decision Making with Pairwise Comparisons. | 7 |
| | | | |
| 4.0 | | **Automated Decision Systems and Expert Systems** | |
| 4 | 4.1 | Basic Concepts of Expert Systems, Applications of Expert Systems, Structure of Expert Systems, Knowledge Engineering, Development of Expert Systems, Location, Based Analytics for Organizations, Cloud Computing in Business | 7 |
| | | | |
| 5.0 | | **Mathematical Models** | |
| 5 | 5.1 | Mathematical Model for Decision Making, Structure of Model, Classes of Models, Regression Models, Simple Linear & Multiple Linear Regression, Validation of Regression Models, Selection of Predictive Variables with Example | 7 |
| | | | |
| 6.0 | | **Classification & Clustering** | |
| 6 | 6.1 | Classification problems, evolution, Classification Tree, Clustering method, partition method, Hierarchical methods, Evaluation of clustering models. | 7 |

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Textbooks:**

1. Carlo Vercellis “Business Intelligence- Data mining & optimization for Decision Making”,Wiley
2. Ramesh Sharda, DursunDelen, Efraim Turban, J. E. Aronson, Ting-Peng Liang, David King, "Business Intelligence and Analytics: System for Decision Support

**References:**

1. Vicki L. Sauter, "Decision Support Systems for Business Intelligence", 2nd Edition,2011, Wiley

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 7 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSVS101 | Natural language Processing | | 04 | - | - | 02 | - |

Evaluation Scheme

| **Course Code** | **Course Name** | **In Semester Evaluation** | | **End Semester Exam (OE/POE)** | | |
|---|---|---|---|---|---|---|
| | | **Term work** | **Min pass** | **Marks** | **Min pass** | **Total (Marks)** |
| PDSVS101 | Natural language Processing | 50 | 40% | - | - | 50 |

**Course Description:** Will be able to implement and evaluate different NLP applications and apply machine learning and deep learning methods for this process. Evaluate various algorithms and approaches for the given task, dataset, and stage of the NLP product.

**Pre-requisites**: Machine learning,paython.

**Course Objectives:**

1. To gain knowledge on NLP.
2. To deal with morphological processing, syntactic parsing, information extraction.
3. To understand probabilistic NLP and classification of text using Pythons NLTK Library.

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Construct** Python programs to manipulate and analyze language data.

**CO2** **Evaluate** key concepts from NLP and linguistics to describe and analyze language.

**CO3** **Write** the data structures and algorithms that are used in NLP.

**CO4** **Differentiate** texts using machine learning and deep learning.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

## Experiment List

1. Write a Python script to process a given text file by tokenizing it into words, removing stop words, and counting the frequency of each word.

2. Use the NLTK library to access a text corpus, create a conditional frequency distribution, and explore lexical resources like WordNet.

3. Implement a Python program that uses regular expressions to identify and extract specific patterns from raw text, such as email addresses and phone numbers.

4. Write a Python script to normalize text (convert to lowercase, remove punctuation) and tokenize it into sentences and words using NLTK.

5. Develop a Python script to access and read text data from web pages and local disk files, then perform basic text processing.

6. Write a Python program to handle and process text containing Unicode characters, including normalization and encoding/decoding.

7. Develop a simple text classification model using the Naive Bayes algorithm to classify text data into predefined categories

8. Implement a basic RNN using a deep learning library (such as Keras) to classify text data.

9. Create a supervised classification model, evaluate its performance using metrics like accuracy and F1-score, and visualize the results.

10. Use an NER tool (like SpaCy) to identify and extract named entities from a given text document.

11. Develop a simple chunker using NLTK to identify and extract phrases (like noun phrases) from sentences in a text.

12. Implement a sentiment analysis model to classify text data (such as movie reviews) into positive and negative sentiments.

13. Use a text-to-speech library (like gTTS) to convert a given text document into speech.

**Text Books**

1. "Speech and Language Processing" by Daniel Jurafsky and James H. Martin

2. "Foundations of Statistical Natural Language Processing" by Christopher D. Manning and Hinrich Schütze

**Reference:**

1. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit" by Steven Bird, Ewan Klein, and Edward Loper

2 "Natural Language Processing: A Concise Introduction" by Jacob Eisenstein

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

**Evaluation Scheme:**

1TERM WORK assessment must be based on the overall performance of the student with every assignment graded from time to time.

2. The grades will be converted to marks as per 'credit and grading system' manual and should be added and averaged.

3. based on above scheme grading and TERM WORK assessment should be done.

.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| PDSPC104 | Open Source tools for Machine Learning & Data Visualization Lab | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| | | | 04 | - | - | 02 | - |

Evaluation Scheme

| **Course Code** | **Course Name** | **In Semester Evaluation** | | **End Semester Exam (OE/POE)** | | |
|---|---|---|---|---|---|---|
| PDSPC104 | Lab I- Open Source tools for Machine Learning & Data Visualization Lab | **Term work** | **Min pass** | **Marks** | **Min pass** | **Total (Marks)** |
| | | 25 | 40% | 25 | 40 | 50 |

**Course Description:** The course will start with a discussion of how machine learning is different than descriptive statistics, and introduce the scikit learn toolkit through a tutorial. The issue of dimensionality of data will be discussed, and the task of clustering data, as well as evaluating those clusters, will be tackled.

**Pre-requisites**: Python/R Programming.

**Course Objectives:**

1. To define the foundational concepts and terminology of machine learning.
2. To Identify different types of machine learning tasks
3. To Describe ethical considerations and biases in machine learning models and data
4. To discuss the implications of model selection and evaluation metrics on decision-making.

**Course Outcomes:** after the end of this course students will be able to

CO1 **Construct** the FIND-S algorithm and its purpose in machine learning

CO2 **Design** linear regression using Python to analyze relationships between variables.

CO3 **Evaluate** the results of linear regression models on real-world datasets

CO4 **Examine** the effectiveness of the back-propagation algorithm for finite words classification..

## Experiment List:

1. Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.

2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.

2. Extract data from a database using Python/R.

3. Implement linear regression using Python/R.

4. Implement logistic regression using python/R.

5. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.

6. Implement Random Forest classification in Python/R.

7. Implement the finite words classification system using the Back-propagation algorithm.

8. Implement k-nearest neighbor classification using Python/R.

9. Implement Naive Bayes theorem to classify English text.

10. Tokenization and Word Frequency Counting

11. Regular Expressions for Pattern Detection

13. Text Normalization and Tokenization

**Text Books**

1. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron

2. "Machine Learning: A Probabilistic Perspective" by Kevin P. Murphy

**Reference:**

1. Python Machine Learning" by Sebastian Raschka and Vahid Mirjalili

2 . "Introduction to Machine Learning with Python: A Guide for Data Scientists" by Andreas C. Müller and Sarah Guido

- Internal Assessment (T1, T2 and FET):

1.T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2.Fifth module (MCQ quiz) will be assessed for 5 marks separately.

**Evaluation Scheme:**

1TERM WORK assessment must be based on the overall performance of the student with every assignment graded from time to time.

2.The grades will be converted to marks as per 'credit and grading system' manual and should be added and averaged.

3.Based on above scheme grading and TERM WORK assessment should be done.

.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPC201 | Statistics For Signal Analysis | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPC201 | Statistics For Signal Analysis | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** This course focuses on the problem of supervised and unsupervised learning from the perspective of modern statistical learning theory, starting with the theory of multivariate function approximation from sparse data. It develops basic tools such as regularization, including support vector machines for regression and classification

**Pre-requisites**: Linear Algebra and Multivariate calculus

**Course Objectives:**

1. To provide an overview of the theories and current practices, required by students who intend to specialize in this field
2. To solve the complex problems in Machine learning applications in different fields.
3. To uncover the common statistical principles underlying this diverse array of techniques.

**Course Outcomes:** after the end of this course students will be able to

CO1 **Evaluate** learning methods statistically and recommend the optimal one for applications.

CO2 **Create** the modern statistical techniques with statistical software such as R.

CO3 **Explain** global and local methods for prediction

CO4 **Compose** the classification, clustering and dimension reduction techniques.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Fundamentals of Probability and Statistical Learning Theory** | |
| 1 | 1.1 | Discrete Distributions, Random Variables, BinomialDistributions, Geometric Distributions, Plug- in estimators, empirical risk minimization, Linear threshold functions, perceptron algorithm, Variance and bias trade-off. Risk bounds, Concentration inequalities, Uniform convergence, Rademacher averages; combinatorial dimensions | 9 |
| | | | |
| 2.0 | | **Regression Techniques and Minimax Strategies** | |
| 2 | 2.1 | Minimax strategies for log loss, linear loss, and quadratic loss Linear regression, Subset selection, Ridge regression, Lasso, Principal components regression. | 9 |
| | | | |
| 3.0 | | **Regression** | |
| 3 | 3.1 | Polynomial regression, Splines, Local smoothers. K means, Spectral clustering, Clustering evaluation, Model-based clustering, EM algorithm, Latent Dirichlet Allocation model. | 9 |
| | | | |
| 4.0 | | **Classification Techniques** | |
| 4 | 4.1 | Classification theory and evaluation, Naïve Bayes classifiers, Logistic Regression with regularization, Linear discriminant analysis. | 9 |
| | 4.2 | Regression tree, Pruning, Tree ensembles, AdaBoost, AdaBoost as I-projection, Convergence and consistency of AdaBoost | |
| | | | |
| 5.0 | | **Support Vector Machines and Kernel Methods** | |
| 5 | 5.1 | Support Vector Machine Maximal Margin Classifier, Support Vector Classifiers, Support Vector Machines with more than Two Classes, kernel Methods | 9 |

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Text Books**

1 George F Luger, *Artificial Intelligence*, Pearson Education Publications, 2001.
2 Elaine Rich and Knight, *Artificial Intelligence*, Mcgraw-Hill Publications, 2017

**Reference:**

1. David Rolston, *Principles of AI and Expert system development,* MGH, 1988.

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPC202 | Optimization Techniques for Analytics | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPC202 | Optimization Techniques for Analytics | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** Optimization is the search for the best and most effective solution. In this course, optimization techniques will be examined through a Business Analytics. The students will be introduced to the theory, algorithms, and applications of optimization. Linear and integer programming will be applied to problems involving data.

**Pre-requisites**: Linear Algebra and Numerical Methods

**Course Objectives:**

1. To introduce optimization techniques using both linear and non-linear programming.
2. To enable the students to frame engineering minima maxima problems in the framework of optimization problems.

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Evaluate** the effectiveness of different optimization techniques in achieving engineering objectives and solving complex problems

**CO2** **Identify** key concepts and components of linear programming.

**CO3** **Analyze** effectiveness and performance of quadratic programming solutions.

**CO4** **Explain** Integer Programming and Optimization Methods

**CO5** **Evaluate** the impact of Recent Trends in Optimization

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Engineering application:** | |
| 1 | 1.1 | Engineering application of Optimization, Formulation of design problems as mathematical programming problems, General Structure of Optimization Algorithms | 9 |
| | | | |
| 2.0 | | **Linear & Nonlinear Programming:** | |
| 2 | 2.1 | Linear Programming, modeling (selected models and as a modeling tool), simplex method and modeling language (AMPL). network flow problems | 9 |
| | 2.2 | Nonlinear Programming, applied models (logistics egression/classification), unconstrained problems, convex optimization, inequality constraints | |
| | | | |
| 3.0 | | **Quadratic Programming:** | |
| 3 | 3.1 | Quadratic Programming, applied models,(least-squares regression, portfolio selection, support vector machines) | 9 |
| | | | |
| 4.0 | | **Integer Programming:** | |
| 4 | 4.1 | Integer Programming, selected problem classes (knapsack,set covering/partition/packing, logical relations), basic branch and bound and cutting planes methods | 9 |
| | | | |
| 5.0 | | **Recent trends:** | |
| 5 | 5.1 | Recent trends: Applications of ant colony optimization, genetics and linear and quadratic programming in real world applications. | 9 |

**Text Books**

1 Laurence A. Wolsey (1998). *Integer programming*. Wiley.
2 John K. Karlof (2006). *Integer programming: theory and practice*.CRC Press.

**Reference:**

1. Dimitris Bertsimas; Robert Weismantel, *Optimization over integers. Dynamic Ideas*.

2. Der-San Chen; Robert G. Batson; Yu Dang, *Applied Integer Programming: Modeling and Solution,* John Wiley and Sons.

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPC203 | Deep Learning | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPC203 | Deep Learning | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** This course is an introduction to deep learning, a branch of machine learning concerned with the development and application of modern neural networks. Deep learning algorithms extract layered high- level representations of data in a way that maximizes performance on a given task.

**Pre-requisites**: Knowledge in Linear Algebra and Probability, Machine learning

**Course Objectives:**

1. To provide exposure to deep learning topics
2. To discuss deep learning methods for working with sequential data
3. To illustrate how to apply such deep learning mechanisms to various learning problems

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Explain** machine learning algorithms to simple datasets, demonstrating data preprocessing, model training, and evaluation.

**CO2** **Develop** different activation functions to neural network layers

**CO3** **Build** application using Convolutional Neural Networks & Sequence modeling

**CO4** **Evaluate** the effectiveness of RNNs for different types of sequential tasks.

**CO5** **Design** a problem-solving strategy using Advanced Search and Optimization Techniques

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Introduction** | |
| 1 | 1.1 | Overview of machine learning, linear classifiers, loss functions Introduction to Tensor Flow: Computational Graph, Key highlights, Creating a Graph, Regression example, Gradient Descent, Tensor Board, Modularity, Sharing Variables, Keras. | 9 |
| | | | |
| 2.0 | | **Activation Functions** | |
| 2 | 2.1 | Sigmoid, ReLU, Hyperbolic Fns, Soft max Perceptrons: What is a Perceptron, XOR Gate Artificial Neural Networks: Introduction, Perceptron Training Rule, Gradient Descent Rule, vanishing gradient problem and solution. | 9 |
| | | | |
| 3.0 | | **Convolutional Neural Networks** | |
| 3 | 3.1 | Introduction to CNNs, Kernel filter, Principles behind CNNs, Multiple Filters, problem and solution of under fitting and over fitting | 9 |
| | | | |
| 4.0 | | **Recurrent Neural Networks** | |
| 4 | 4.1 | Introduction to RNNs, Unfolded RNNs, Seq2Seq RNNs, LSTM, GRU, Encoder Decoder architectures | 9 |
| | | | |
| 5.0 | | **Deep Learning applications** | |
| 5 | 5.1 | Image segmentation, Object detection, Attention model for computer vision tasks, Natural Language Processing, Speech Recognition, Video Analytics | 9 |

**Text Books**

1. Good fellow, I., Bengio,Y., and Courville, A., Deep Learning, MIT Press, 2016. .
2. Josh Patterson, Adam Gibson, Deep Learning: A Practitioner's Approach, OReilly, 2017.

**Reference:**

1. Bishop, C., M., Pattern Recognition and Machine Learning, Springer, 2006.

2. Yegnanarayana, B., Artificial Neural Networks PHI Learning Pvt. Ltd, 2009.

3. Golub, G., H., and Van Loan, C.,F., Matrix Computations, JHU Press,2013.

4. Satish Kumar, Neural Networks: A Classroom Approach, Tata McGraw-Hill Education, 2004.

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| PDSPE201 | Program Specific Elective II: Bioinformatics | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| | | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| PDSPE201 | Program Specific Elective II: Bioinformatics | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| | | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** The course aims to give overview of concepts and databases from bioinformatics and use of artificial intelligence in biology and medical field.

**Pre-requisites**: Artificial Intelligence.

**Course Objectives:**

1. To discuss techniques to search and retrieve information from genomic and proteomic databases.
2. To illustrate how to compare and analyze biological sequences and how to interpret the results of their analyses
3. To familiarize students with applications of AI techniques for biological and medical problems

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Choose** biological data, submission and retrieval it from databases and design databases to store the information.

**CO2** **Experiment** the most important bioinformatics databases, perform text- and sequence based searches, and analyze the results in light of molecular biological knowledge.

**CO3** **Identify** machine learning techniques in bioinformatics

**CO4** **Evaluate** the impact of different preprocessing techniques

**CO5** **Develop** case studies for applying medical image processing and supervised machine learning techniques in bioinformatics

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Basics of Bioinformatics**: | |
| 1 | 1.1 | Introduction to Bioinformatics; Computers in Biology to understand Biological System; Open resources in Bioinformatics. | 9 |
| | | | |
| 2.0 | | **Sequence Analysis** | |
| 2 | 2.1 | Biological background for sequence analysis; Sequence alignment: Global, Local, Pairwise and Multiple sequence analysis; Algorithm for alignments; Database Searching; Tools for Sequence alignment. | 9 |
| | | | |
| 3.0 | | **Biological Databases**: | |
| 3 | 3.1 | Database concepts; Introduction to Data types and source; Protein Sequence and Structural Databases; Nucleic acid databases; Genome databases; Specialized Databases; Carbohydrate Databases; Clinically relevant drug-drug interactions databases; Information retrieval from Biological databases: Entrez system, TCGA data bases, Bioportal | 9 |
| | | | |
| 4.0 | | **Feature Interpretation for Biological Learning:** | |
| 4 | 4.1 | Introduction, Normalization Techniques for Gene Expression Analysis, Data preprocessing of Mass Spectrometry Data, Techniques for MS Data **Analysis**, Data Preprocessing for Genomic Sequence Data, Ontologies in Bioinformatics | 9 |
| | | | |
| 5.0 | | **Medical Image Processing**: | |
| 5 | 5.1 | Feature extraction, identifying positions of features - Normalization data cleaning and transformation, deep learning for medical image processing | 9 |

| | | | |
|---|---|---|---|
| | 5.2 | Recent case studies and applications of supervised learning techniques in Bioinformatics | |

**Text Books**

1.Lesk, A.M., *Introduction to Bioinformatics*; Oxford University Press, UK, Fourth edition, 2014

2.Gretchen Kenney, *Bioinformatics: Principles and Analysis*; Syra wood Publishing House USA, 2016

**Reference:**

1. Hooman Rashidi, Lukas K., *Bioinformatics Basics: Applications in Biological Science and Medicine*, Buehler Publisher: CRC Press/Taylor & Francis Group, 2005.
2. Ian Goodfellow, Deep Learning (Adaptive Computation And Machine Learning Series) , The MIT Press , 2016

- Internal Assessment (T1, T2 and FET):

1.T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2.Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1.Question paper will comprise of 5 questions, each carrying 10 marks.

2.The duration of end semester examination shall be 2 hours.

3.The students need to solve all 5 questions.

4.Question No.1 will be compulsory and based on entire syllabus.

5.Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PDSPE202 | Program Specific Elective II: Pattern Recognition | 03 | - | - | 03 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PDSPE202 | Program Specific Elective II: Pattern Recognition | 10 | 10 | 5 | 25 | 40% | 50 | 40% | 75 |

**Course Description:** This course focuses on the underlying principles of pattern recognition and on the methods of machine intelligence used to develop and deploy pattern recognition applications in the real world. Emphasis is placed on the pattern recognition application development process, which includes problem identification, concept development, algorithm selection, system integration, test and validation the design analysis, and development of methods for the classification or description of patterns, objects.

**Pre-requisites**: Linear Algebra, Probability and Statistics.

**Course Objectives:**

1. To Learn the fundamentals of pattern recognition and its relevance to classical and modern problems.
2. To Identify where, when and how pattern recognition can be applied.
3. To Introduction to more recent applications of pattern recognition, such as cognitive neuroscience and bioinformatics

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Justify** the general working of the Machine Learning algorithms..

**CO2** **Create** a comprehensive MLE-based estimation approach

**CO3** **Evaluate** the effectiveness of statistical Pattern Recognization.

**CO4** **Analyze** the concept of VC-dimension and its significance in understanding model capacity and generalization.

**CO5** **Explain** effectiveness of Advanced Artificial Neural Networks

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| 1.0 | | **Introduction** | |
| 1 | 1.1 | Introduction to Statistical Pattern Recognition Overview of Pattern Classifiers Bayes Classifier and Nearest Neighbor Classifier Minmax and Neymann-Pearson Classifiers. | 9 |
| | | | |
| 2.0 | | **Statistical Estimation Methods** | |
| 2 | 2.1 | Maximum Likelihood Estimation of various class conditional densities Bayesian estimation and MAP estimates Exponential family of densities Sufficient Statistics and recursive formulation of ML and Bayesian estimates. | 9 |
| | | | |
| 3.0 | | **SYNTACTIC PATTERN RECOGNITION:** | |
| 3 | 3.1 | Overview of Syntactic Pattern Recognition, Grammar Based Approaches and Applications, Examples of String Generation as Pattern escription,2-D line Drawing Description Grammar, Character Description using PDL, Object Description using Projected Cylinder Models, Block World Description Models, Heuristic Generation of Grammars. | 9 |
| | | | |
| 4.0 | | **PAC Learning, Generalization, and VC-Dimension** | |
| 4 | 4.1 | PAC learning framework Empirical Risk Minimization Concept of Generalization Bias-Variance Trade-off Thorough understanding of VC-dimension with examples. | 9 |
| | | | |
| 5.0 | | **Artificial Neural Networks: Architecture, Training, and Applications** | |
| 5 | 5.1 | Artificial Neural Network and related terms Activation functions Feedforward Neural Network with Backpropagation algorithm Regularization and momentum in Artificial Neural Network | 9 |

**Text Books**

1. Pattern Classification', R. O. Duda, P.E. Hart and D.G. Stork, Johy Wiley, Publication
2. 'Pattern Recognition and Machine Learning', C.M. Bishop, Springer Publication

**Reference:**

1. Deep Learning, I. Goodfellow, Y. Bengio and A. Courville C. M. Bishop, MIT Press Publication.
2. Learning From Data, Yaser S. Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tien Lin, AML Book

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules and T2 (Test 2) should be based on next two modules, for 10 marks each.

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

1. Question paper will comprise of 5 questions, each carrying 10 marks.
2. The duration of end semester examination shall be 2 hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| PSTRM201 | Research Publications and Ethics | 02 | - | - | 02 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| PSTRM201 | Research Publications and Ethics | | 10 | 5 | 15 | 40% | 35 | 40% | 50 |

**Course Description**

This course will provide an opportunity for participants to establish or advance their understanding of research through critical exploration of research language, ethics, and approaches.

**Course Objectives**

- To know research ethics.
- To understand research approaches and their applications.

**Course Outcomes:** After the successful completion of the course students will able to:

- **CO1** **Identify** the ethical standards in research
- **CO2** **Analyse** research literature review and identify gaps in literature pertaining to a particular research issue.
- **CO3** **Design** and conduct research
- **CO4** **Describe** the outcomes of research by preparing a paper for publication or patenting

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

**Course Contents**

| Module | Unit | Description | Hours |
|---|---|---|---|
| **1.0** | | **Publication ethics:** | **9** |
| 1 | 1.1 | definition, introduction, and importance - best practices / standards setting initiatives and guidelines: COPE, WAME, etc. - Conflicts of interest - | |
| | 1.2 | Publication misconduct: definition, concept, problems that lead to unethical behavior and vice versa, types - Violation of publication ethics, authorship, and contributor ship - Identification of publication misconduct, complaints, and appeals - Predatory publisher and journals. | |
| **2.0** | | **Ethics with respect to science and research -** | **9** |
| 2 | 2.1 | Intellectual honesty and research integrity - Scientific misconducts: Falsification, Fabrication and Plagiarism (FFP) - Redundant Publications: duplicate and overlapping publications, | |
| | 2.2 | salami slicing - Selective reporting and misrepresentation of data. | |
| | | | |
| **3.0** | | **Open access publications and initiatives -** | **9** |
| 3 | 3.1 | SHERPA/RoMEO online resource to check publisher copyright & self-archiving policies - Software tool to identify predatory publications developed by SPPU - Journal finger / journal suggestion tools viz. JANE, Elsevier Journal Finder, Springer, Journal Suggester, etc. | |
| | | | |
| **4.0** | | **Databases:** | **9** |
| 4 | 4.1 | Indexing databases, Citation databases: Web of Science, Scopus, etc. Research Metrics: Impact Factor of journal as per Journal Citations Report, SNIP, SJR, IPP,. | |
| | 4.2 | Cite Score - Metrics: h-index, g index, i10 Index, altmetrics | |

**Text Books**

1. Gupta S. C. and Kapoor V. K, Fundamentals of Mathematical Statistics, Sultan Chand & Company New Delhi.
2. Creswell, J. W. Research design: Qualitative, quantitative, and mixed methods approach. 5th Ed. Thousand Oaks, CA: Sage, 2018.

**References**

1. Nicholas H. Steneck. Introduction to the Responsible Conduct of Research. Office of Research Integrity. 2007. Available at: https://ori.hhs.gov/sites/default/files/rcrintro.pdf
2. Anderson B.H., Dursaton, and Poole M.: Thesis and assignment writing, Wiley Eastern 1997.
3. Gupta S. C. and Kapoor V. K, "Fundamentals of Mathematical Statistics", Sultan Chand & Company New Delhi.

**Internal Assessment (T1, T2 and FET)**

1. T1 (Test 1) should be based on first two modules for 10 marks.
2. FET shall be assessed for 5 marks separately.

**End Semester Examination**

1. Question paper will be of 35 marks comprise of 5 questions, each carrying 07 marks.
2. The duration of end semester examination shall be Two hours.
3. The students need to solve all 5 questions.
4. Question No.1 will be compulsory and based on entire syllabus.
5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| | | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| P70OE2XX | Open/Multidisciplinary Elective II-Big Data Analytics | 02 | - | - | 02 | - | - |

Evaluation Scheme

| Course Code | Course Name | Evaluation Scheme (In Semester) | | | | | End Semester Exam (ESE) | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T1 | T2 | FET | Total | Min pass | Marks | Min pass | Total (Marks) |
| P70OE2XX | Open/Multidisciplinary Elective II-Big Data Analytics | | 10 | 5 | 15 | 40% | 35 | 40% | 50 |

**Course Description**:

This course gives an overview of Big Data, i.e. storage, retrieval and processing of big data. In addition, it also focuses on the "technologies", i.e., the tools/algorithms that are available for storage, processing of Big Data. It also helps a student to perform a variety of "analytics" on different data sets and to arrive at positive conclusions.

**Pre-requisites:** Should have knowledge of one Programming Language (Java preferably), Practice of SQL (queries and sub queries), exposure to Linux Environment.

**Course Objective:**

1. To develop a solid understanding of the fundamental concepts, principles, and technologies related to big data analytics, including data preprocessing, storage, processing, and analysis.

2. To Learn various big data platforms and frameworks, such as Apache Hadoop and distributed file systems, and understand how they facilitate the storage and processing of massive datasets.

3. To Gain proficiency in handling and processing diverse types of big data, including structured, semi-structured, and unstructured data, using appropriate tools and techniques.

4. To Gain hands-on experience with big data analytics tools and technologies through practical exercises, lab sessions, and projects.

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Analyze** Big Data and its Business Implications.

**CO2** **Identify** the components of HDFS

**CO3** **Design** Job Execution in MapReduce

**CO4** **Examine** various data processing Using Pig

**CO5** **Explain** Machine Learning Techniques using R.

| Course Contents | | | |
|---|---|---|---|
| **Module** | **Unit** | **Description** | **Hours** |
| **1.0** | | **Introduction to Big Data and Hadoop:** | |
| 1 | 1.1 | Types of Digital Data, Introduction to Big Data, Big Data Analytics, History of Hadoop, Apache Hadoop, Analyzing Data with Unix tools, Analyzing Data with Hadoop | 7 |
| | 1.2 | Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Infosphere BigInsights and Big Sheets. | |
| | | | |
| **2.0** | | **Hdfs (Hadoop Distributed File System):** | |
| 2 | 2.1 | The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow | 7 |
| | 2.2 | Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures. | |
| | | | |
| **3.0** | | **Map Reduce:** | |
| 3 | 3.1 | Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, | 7 |
| | 3.2 | Map Reduce Types and Formats, Map Reduce Features. | |
| | | | |
| **4.0** | | **Hadoop Eco System:** | |
| 4 | 4.1 | Pig: Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. | 7 |
| | | | |
| **5.0** | | **Hadoop Eco System:** | |
| 5 | 5.1 | Hive: Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions | 7 |
| | 5.2 | Machine Learning: Introduction, Supervised Learning, Unsupervised Learning, Collaborative Filtering. Big Data Analytics with BigR. | |
| | | | |

**Textbooks:**

1. Tom White “Hadoop: The Definitive Guide” Third Edit on, O’reily Media, 2012.

2. Seema Acharya, Subhasini Chellappan, "Big Data Analytics" Wiley 2015.

**References:**

1. Paul Zikopoulos ,Dirk DeRoos , Krishnan Parasuraman , Thomas Deutsch , James Giles , David Corigan , "Harness the Power of Big Data The IBM Big Data Platform ", Tata McGraw Hill Publications, 2012.

2. Bill Franks, “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics”, John Wiley & sons, 2012.

3. Pete Warden, “Big Data Glossary”, O’Reily, 2011

- Internal Assessment (T1, T2 and FET):

1. T1 (Test 1) should be based on first two modules

2. Fifth module (MCQ quiz) will be assessed for 5 marks separately.

- End Semester Examination:

  1. Question paper will comprise of 5 questions, each carrying 7 marks.
  2. The duration of end semester examination shall be 2 hours.
  3. The students need to solve all 5 questions.
  4. Question No.1 will be compulsory and based on entire syllabus.
  5. Remaining question (Q.2 to Q.5) will be selected from all the modules.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| PDSVS201 | Web & Social Media Analytics | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| | | | 04 | - | - | 02 | - |

Evaluation Scheme

| **Course Code** | **Course Name** | **In Semester Evaluation** | | **End Semester Exam (OE/POE)** | | |
|---|---|---|---|---|---|---|
| PDSVS201 | Web & Social Media Analytics | **Term work** | **Min pass** | **Marks** | **Min pass** | **Total (Marks)** |
| | | 50 | 40% | - | - | 50 |

**Course Description:** This course provides hands-on experience in applying Natural Language Processing (NLP) techniques and analytics tools to extract insights from text data, analyze customer feedback, optimize web content, and perform web scraping. Students will gain practical skills in preprocessing text, designing decision support systems, sentiment analysis, web analytics, and social network analysis.

**Pre-requisites**: NLP.

**Course Objectives:**

1. To Knowledge on decision support systems.
2. To Apply natural language processing concepts on text analytics.
3. To Understand sentiment analysis.
4. To Knowledge on search engine optimization and web analytics

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Explain** and classify various NLP techniques including stop word elimination, stemming, lemmatization, POS tagging, and lexical analysis.

**CO2** **Identify** techniques for extracting meaningful information from text corpora.

**CO3** **Summarize** the components and functionalities of a basic Decision Support System (DSS).

**CO4** **Analyze** sentiment trends and patterns in customer reviews and social media feedback.

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

## Experiment List

1. Pre-processing text document using NLTK of Python
    a. Stop word elimination
    b. Stemming
    c. Lemmatization
    d. POS tagging
    e. Lexical analysis
2. Design and implement a basic Decision Support System.
3. Sentiment analysis on customer review on products
4. Extract meaningful information from a corpus of text documents.
5. Analyzing Customer Feedback on Social Media
6. Speech Sentiment Analysis
7. Web analytics
    a. Web usage data (web server log data, click stream analysis)
    b. Hyperlink data
8. Search engine optimization- implements spamdexing
9. Use Google analytics tools to implement the following
    a. Conversion Statistics
    b. Visitor Profiles
10. Use Google analytics tools to implement the Traffic Sources.
11. Social Network Analysis
12. Optimization and What-If Analysis
13. Web Scraping for Product Price Comparison

**Text Books**

1. Ramesh Sharda, DursunDelen, Efraim Turban, BUSINESS INTELLIGENCE AND ANALYTICS: SYSTEMS FOR DECISION SUPPORT, Pearson Education.

**Reference:**

1. RajivSabherwal, Irma Becerra- Fernandez," Business Intelligence –Practice, Technologies and Management", John Wiley 2011.

2. Lariss T. Moss, Shaku Atre, "Business Intelligence Roadmap", Addison-Wesley It Service.

3. YuliVasiliev, "Oracle Business Intelligence: The Condensed Guide to Analysis and Reporting", SPD Shroff, 2012.

**Evaluation Scheme:**

1TERM WORK assessment must be based on the overall performance of the student with every assignment graded from time to time.

2. The grades will be converted to marks as per 'credit and grading system' manual and should be added and averaged.

3. based on above scheme grading and TERM WORK assessment should be done.

.

****

**Curriculum Structure of M. Tech. Artificial Intelligence & Data Science**
**(AY: 2024-25)**

| Course Code | Course Name | Teaching Scheme (Hr/week) | | | Credits Assigned | | |
|---|---|---|---|---|---|---|---|
| PDSPC204 | Advanced Data Visualization Lab | Theory | Practical | Tutorial | Theory | Practical | Tutorial |
| | | | 04 | - | - | 02 | - |

Evaluation Scheme

| **Course Code** | **Course Name** | **In Semester Evaluation** | | **End Semester Exam (OE/POE)** | | |
|---|---|---|---|---|---|---|
| PDSPC204 | Advanced Data Visualization Lab | **Term work** | **Min pass** | **Marks** | **Min pass** | **Total (Marks)** |
| | | 25 | 40% | 25 | 40% | 50 |

**Course Description:** Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. It is the representation of information and data through use of common graphics, such as charts, plots, info graphics, and animations.

**Pre-requisites**: NLP.

**Course Objectives:**

1. To Analyze and solve real time data visualization scenarios using Python/R integrating with Tableau.

**Course Outcomes:** after the end of this course students will be able to

**CO1** **Integrate** with Tableau for various data visualization scenarios..

**CO2** **Design** visualization dashboard to support the decision making on large scale data..

**CO3** **Demonstrate** the analysis of large dataset using various visualization techniques and tool

## Experiment List

1. Acquire a dataset and create various plots (line, bar, scatter) to visualize the data.

2.Perform statistical analysis on a dataset to calculate mean, median, mode, variance, and standard deviation.

3.Apply K-means and hierarchical clustering on a dataset and visualize the clusters.

4.Conduct correlation, regression, and ANOVA on a dataset to explore variable relationships and visualize results

5Analyze a financial dataset using clustering, histograms, and heat maps to identify patterns and trends.

6. Perform time-series analysis on stock market data to identify trends and forecast future prices.

7. Visualize large datasets from healthcare, census, and geospatial domains to highlight key insights.

8. Visualize real-time streaming data from stock market and weather forecasts, providing dynamic updates.

9. Analyze market-basket data to identify association rules and visualize purchasing patterns.

10. Perform text analysis on web analytics data using word clouds, sentiment analysis, and topic modeling.

11. Create interactive visualizations using tools like Plotly or Bokeh. Enable users to explore the data by interacting with the visual elements.

12. Apply dimensionality reduction techniques like PCA (Principal Component Analysis) and t-SNE (t-Distributed Stochastic Neighbor Embedding) on a high-dimensional dataset and visualize the reduced dimensions

**Text Books**

1. Tamara Munzer, Visualization Analysis and Design, 1st edition, CRC Press, United States, 2015.

2. Michael Fry, Jeffrey Ohlmann, Jeffrey Camm, James Cochran, Data Visualization: Exploring and Explaining
with Data, South-Western College Publishing, 2021

**Reference:**

1.Dr. Chun-hauh Chen, W. K. Hardle, A. Unwin, Handbook of Data Visualization, 1st edition, Springer publication, Germany, 2008.

2. Ben Fry, Visualizing Data, 1st edition, O'Reilly Media, United States, 2008.

3. Avril, A little book of R for multivariate analysis, 1st edition, Welcome Trust Sanger Institute, United Kingdom, 2013

**Evaluation Scheme:**

1TERM WORK assessment must be based on the overall performance of the student with every assignment graded from time to time.

2. The grades will be converted to marks as per 'credit and grading system' manual and should be added and averaged.

3. based on above scheme grading and TERM WORK assessment should be done.

.

****