ST. JOSEPH'S
COLLEGE OF ENGINEERING
AND TECHNOLOGY,
- PALAI -
AUTONOMOUS

Choondacherry P.O., Pala, Kottayam - 686579
Kerala, India

SYLLABUS

Minor *in*

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

OFFERED BY: Department of Artificial Intelligence and Data Science (AD) ELIGIBLE DEPARTMENTS: CE, EC, ER, EE, ME

2024 SCHEME

Semester IV, V & VI

Signed By: Devassia V P
Reason: Approved by SJCET Academic Council held on 25.11.2025

CURRICULUM

Minor in Artificial Intelligence and Data Science											
Sl. No:	**Semester**	**Course Code**	**Course Title (Course Name)**	**Credit Structure**			**SS**	**Total Marks**		**Credits**	**Hrs./ Week**
				L	**T**	**P**		**CIE**	**ESE**		
1	3	24SJMNADT301	Introduction to Artificial Intelligence*/MOOC#	3	1	0	5	40	60	4	4
2	4	24SJMNADT401	Concepts in Machine Learning*/MOOC#	3	1	0	5	40	60	4	4
3	5	24SJMNADT501	Deep Learning*/MOOC#	3	1	0	5	40	60	4	4
4	6	24SJMNADT601	Natural Language Processing*/MOOC#	3	0	0	5	40	60	3	3
Total							20			15	15

* Students must register for theory courses listed in the 3rd and 4th semesters of the Minor curriculum.

Students who fail a theory course listed in the Minor curriculum are permitted to register for an alternate MOOC course specified in the Minor curriculum.

SEMESTER S4

CONCEPTS IN MACHINE LEARNING

Course Code	**24SJMNADT401**	**CIE Marks**	40
Teaching Hours/Week (L: T: P: R)	3:1:0:0	**ESE Marks**	60
Credits	4	**Exam Hours**	2 Hrs. 30 Min.
Prerequisites (if any)	Basic understanding of probability theory and linear algebra	**Course Type**	Theory

Course Objectives:

1. To understand the basic concepts of data science.
2. To understand the fundamental concepts and algorithms in machine learning.
3. To provide machine learning based solutions to real world problems.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Introduction to Data Science:** A brief introduction to data – structured, unstructured, semi-structured, data sets & patterns, Brief history of Data Science, Introduction to Data Science, Importance of Data Science, Differences between AI, ML, DL, Data Science & Data Analytics, Real world applications of data science, Steps in data science process. Tools and Skills Needed – brief introduction of platforms, tools, frameworks, languages, databases and libraries. **Overview of machine learning:** Machine learning paradigms-supervised, semi-supervised, unsupervised, reinforcement learning. **Supervised learning**: Input representation, Hypothesis class, Version space, Vapnik-Chervonenk is (VC) Dimension, Noise, Learning Multiple classes, Model Selection and Generalization.	**11**

2	**Supervised Learning and Parameter Estimation:** **Regression** - Linear regression with one variable, Linear regression with multiple variables, solution using gradient descent algorithm and matrix method, basic idea of overfitting in regression. **Linear Methods for Classification-** Logistic regression, Naive Bayes, Decision tree algorithm ID3. **Basics of parameter estimation** - Maximum Likelihood Estimation (MLE) and Maximum a Posteriori estimation (MAP). Bias-Variance decomposition.	**11**
3	**Support Vector Machines** - Introduction, Maximum Margin hyperplanes, Mathematics behind Maximum Margin Classification, soft margin SVM classifier, non-linear SVM, Kernels for learning non-linear functions, polynomial kernel, Radial Basis Function (RBF), Kernel Trick. **Classification Assessment & SVM Classifier** Classification Performance measures - Precision, Recall, Accuracy, F-Measure, Receiver Operating Characteristic Curve (ROC), Area Under Curve AUC. Bootstrapping, Cross Validation.	**11**
4	**Neural Networks (NN) AND Unsupervised Learning** Perceptron, Neural Network - Multilayer feed forward network, Activation functions (Sigmoid, ReLU, Tanh), Back propagation algorithm. **Ensemble methods**- Voting, Bagging, Boosting. **Clustering** - Similarity measures, Hierarchical Agglomerative Clustering, K-means partitional clustering, Expectation maximization (EM) for soft clustering. **Dimensionality reduction** – Principal Component Analysis, factor Analysis, Multidimensional scaling, Linear Discriminant Analysis.	**11**

Course Assessment Method

(CIE: 40 marks, ESE: 60 marks)

Continuous Internal Evaluation Marks (CIE):

Attendance	Assignment	Internal Examination-1 (Written)	Internal Examination- 2 (Written)	Total
5	**15**	**10**	**10**	**40**

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions

Part A	Part B	Total
• 2 Questions from each module. • Total of 8 Questions, each carrying 3 marks **(8x3 =24 marks)**	• Each question carries 9 marks. • Two questions will be given from each module, out of which 1 question should be answered. • Each question can have a maximum of 3 subdivisions. **(4x9 = 36 marks)**	**60**

Course Outcomes (COs)

At the end of the course students should be able to:

	Course Outcome	Bloom's Knowledge Level (KL)
CO1	Recall the fundamental concepts and applications of data science, and make inferences on key important and Illustrate Machine Learning concepts and basics of supervised learning concepts.	**K3**
CO2	Describe supervised learning concepts (regression, linear classification) and illustrate basics of parameter estimation models.	**K3**
CO3	Evaluate the performance of machine learning models and the working of classifier SVM classifier model.	**K3**
CO4	Illustrate the concepts of Multilayer neural network and unsupervised learning concepts.	**K3**

Note*: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create*

CO-PO Mapping

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1	2	2	3	2	2						
CO2	3	3	3	2	2						
CO3	3	3	3	2							
CO4	3	3	3	2							

Note: *1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High)*

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Introduction to Machine Learning	Ethem Alpaydin	MIT Press	2/e, 2010
2	Machine Learning	Tom Mitchell	McGraw-Hill	1997
3	Fundamentals of Data Science	Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare	CRC press	
4	Data mining Concepts and Techniques	Jiawei Han, Michelin Kamber, Jian Pei	Morgan Kaufmann Publishers	3/e, 2012

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Neural Networks for Pattern Recognition	Christopher Bishop	Oxford University Press	1995
2	Machine Learning: A Probabilistic Perspective	Kevin P. Murphy	MIT Press	2012
3	The Elements of Statistical Learning	Trevor Hastie, Robert Tibshirani, Jerome Friedman	Springer	2/e,2007
4	Elements of Machine Learning	P. Langley	Morgan Kaufmann	1995
5	Data Mining Techniques	Arun K. Pujari	Universities Press	
6	Data Science for Business	Foster Provost, Tom Fawcett	O'Reilly Media	

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1, 2, 3, 4	Introduction to Machine Learning by Prof. Sudeshna Sarkar - IITKGP https://onlinecourses.nptel.ac.in/noc25_cs149
1 -Part 1	Python for Data Science by Prof. Ragunathan Rengasamy-IIT Madras https://onlinecourses.nptel.ac.in/noc23_cs99

SEMESTER S5

DEEP LEARNING

Course Code	**24SJMNADT501**	**CIE Marks**	40
Teaching Hours/Week (L: T: P: R)	3:1:0:0	**ESE Marks**	60
Credits	4	**Exam Hours**	2 Hrs. 30 Min.
Prerequisites (if any)	24SJMNADT401	**Course Type**	Theory

Course Objectives:

1. To provide a foundation in the basic concepts and mathematical intuition of Deep Learning.
2. To explore different neural network architectures and their optimization techniques.
3. To understand the applications of deep learning in vision, text, and other real-world domains.
4. To familiarize students with modern tools and frameworks for implementing deep learning models.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Fundamentals of Neural Networks** Introduction to Artificial Neural Networks – Biological inspiration – Perceptron model – Activation functions (Sigmoid, Tanh, ReLU) – Loss functions (Mean Squared Error, Cross-Entropy) – Feedforward neural networks – Gradient descent and the learning process – Backpropagation algorithm (intuitive explanation).	11
2	**Deep Neural Networks and Optimization** Multilayer neural networks – Weight initialization – Vanishing and exploding gradients (conceptual overview) – Regularization techniques (L1, L2, Dropout) – Batch normalization – Optimization algorithms (SGD, Momentum, RMSProp, Adam – concept and comparison).	11
3	**Convolutional Neural Networks (CNN) & Sequential Models** Concept of convolution and pooling – Filters, feature maps, and parameter sharing – CNN architecture and layers – Activation visualization – Applications in image classification and object detection – Transfer learning and fine-tuning (overview). Sequence modeling – Recurrent Neural Networks (RNN) – Vanishing gradient problem (conceptual) – Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) (concepts and intuition)	11

4	**Autoencoders and Generative models** Applications in natural language processing and time-series analysis. Advanced Concepts and Applications - Overview of autoencoders and generative models (Autoencoder, Variational Autoencoder, GANs – concept only) – Recent trends in deep learning research – Case studies and applications in engineering, healthcare, and IoT – Introduction to deep learning frameworks (TensorFlow, Keras, PyTorch).	**11**

Course Assessment Method

(CIE: 40 marks, ESE: 60 marks)

Continuous Internal Evaluation Marks (CIE):

Attendance	Assignment	Internal Examination-1 (Written)	Internal Examination- 2 (Written)	Total
5	**15**	**10**	**10**	**40**

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions

Part A	Part B	Total
• 2 Questions from each module. • Total of 8 Questions, each carrying 3 marks **(8x3 =24 marks)**	• Each question carries 9 marks. • Two questions will be given from each module, out of which 1 question should be answered. • Each question can have a maximum of 3 subdivisions. **(4x9 = 36 marks)**	**60**

Course Outcomes (COs)

At the end of the course students should be able to:

Course Outcome		Bloom's Knowledge Level (KL)
CO1	Explain the structure, learning process, and regularization techniques involved in deep neural networks.	**K2**
CO2	Describe the architecture and functioning of Convolutional Neural Networks (CNN) and their use in solving classification problems.	**K2**
CO3	Explain the concepts of Recurrent Neural Networks (RNN) and their variants (LSTM, GRU) for handling sequential data.	**K2**
CO4	Summarize the working principles of autoencoders and generative models used in advanced deep learning.	**K2**

Note*: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create*

CO-PO Mapping Table (Mapping of Course Outcomes to Program Outcomes)

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1	3	3	3	3							2
CO2	3	3	3	3	3						2
CO3	3	3	3	3	3						2
CO4	3	3	3	3							2

Note: *1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High)*

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Deep Learning	Goodfellow, I., Bengio, Y., and Courville, A.	MIT Press	1 /e, 2016
2	Neural Networks and Deep Learning	Aggarwal, Charu C	Springer International	1/e, 2018

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Deep Learning, Core Concepts, Methods and Applications	M Gopal	Pearson Education	1/e, 2022
2	Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithm	Nikhil Buduma and Nicholas Locascio	O'Reilly Media	1/e, 2017

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1, 2, 3, 4	https:d21.ai https://nptel.ac.in/courses/106106184 https://cs230.stanford.edu/ https://study.iitm.ac.in/ds/course_pages/BSCS3004.html

SEMESTER S6

NATURAL LANGUAGE PROCESSING

Course Code	**24SJMNADT601**	**CIE Marks**	40
Teaching Hours/Week (L: T: P: R)	3:0:0:0	**ESE Marks**	60
Credits	3	**Exam Hours**	2 Hrs. 30 Min.
Prerequisites (if any)	None	**Course Type**	Theory

Course Objectives:

1. To introduce the comprehensive understanding of Natural Language Processing (NLP)
2. To discuss various parsing techniques and ambiguity resolution
3. To discuss the advanced semantic interpretation and knowledge representation

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Introduction to Natural Language Understanding and Linguistics:** Levels of language analysis, Syntax, Semantics, Pragmatics. Linguistic Background, An Outline of English Syntax, Lexicon, POS Tagging, Word Senses.	**8**
2	**Parsing Techniques and Ambiguity Resolution:** Grammars and Parsing Features, Agreement and Augmented Grammars, Grammars for Natural Language, Parsing methods and Efficient Parsing, Ambiguity Resolution, Statistical Methods. Probabilistic Context Free Grammar.	**8**
3	**Semantic Interpretation and Knowledge Representation in NLP:** Semantics and Logical Form, Linking Syntax and Semantics, Ambiguity Resolution, other Strategies for Semantic Interpretation, Scoping and the Interpretation of Noun Phrases. Knowledge Representation and Reasoning, Local Discourse Context and Reference, Using World Knowledge, Discourse Structure, Defining a Conversational Agent.	**10**
4	**Language Models:** Pre-trained Models-BERT, GPT-2, ELMO, RoBERT Applications and Challenges- Machine Translation, Information Retrieval and Extraction, Sentiment Analysis, Text Categorization and Summarization.	**10**

Course Assessment Method (CIE: 40 marks, ESE: 60 marks)

Continuous Internal Evaluation Marks (CIE):

Attendance	Assignment/ Microproject	Internal Examination-1 (Written)	Internal Examination- 2 (Written)	Total
5	15	10	10	40

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions

Part A	Part B	Total
• 2 Questions from each module. • Total of 8 Questions, each carrying 3 marks **(8x3 =24 marks)**	• Each question carries 9 marks. • Two questions will be given from each module, out of which 1 question should be answered. • Each question can have a maximum of 3 subdivisions. **(4x9 = 36 marks)**	**60**

Course Outcomes (COs)

At the end of the course students should be able to:

Course Outcome		Bloom's Knowledge Level (KL)
CO1	Apply syntax, semantics, and pragmatics in NLP tasks and explain accurate lexicon usage, POS tagging, and word sense disambiguation.	**K3**
CO2	Implement and utilize various grammars and parsing methods, including resolving ambiguities	**K3**
CO3	Link syntax and semantics, interpret noun phrases, and use knowledge representation for discourse management.	**K2**
CO4	Implement the models like BERT, GPT-2, ELMO, and RoBERT for tasks like machine translation, sentiment analysis, and text summarization	**K3**

Note*: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create*

CO-PO Mapping

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1	3	3									3
CO2	3	3	3	3	3						3
CO3	3	3	3	3	3						3
CO4	3	3	3	3	3				3	3	3

Note: *1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High)*

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Speech and Language Processing	D. Jurafsky and J. H. Martin	Prentice Hall India	1/e, 2000
2	Natural Language Understanding	James Allen	Benjamin-Cummings	1/e, 1988

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Introduction to Artificial intelligence	Charniak, Eugene	Addison-Wesley	1/e, 1985
2	Modern Information Retrieval	Ricardo Baeza-Yates and Berthier Ribeiro-Neto	Addison-Wesley	1/e, 1999
3	Natural Language Processing and Information Retrieval,	U. S. Tiwary and Tanveer Siddiqui	Oxford University Press	1/e, 2008

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1, 2	https://onlinecourses.nptel.ac.in/noc24_cs39/preview
4	https://link.springer.com/chapter/10.1007/978-3-031-23190-2_2