Deep Learning and Reinforcement Learning		Semester	
Course Code	**BAI701**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:2:0	SEE Marks	50
Total Hours of Pedagogy	40 hours Theory + 8-10 Lab slots	Total Marks	100
Credits	04	Exam Hours	
Examination nature (SEE)	Theory/practical/Viva-Voce /Term-work/Others		

Course objectives:

- Understand the fundamentals of deep learning.
- Know the theory behind Convolutional Neural Networks, RNN.
- Illustrate the strength and weaknesses of many popular deep learning approaches.
- Introduce major deep learning algorithms, the problem settings, and their applications to solve real world problems

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Black board teaching (Chalk and talk)
2. PPT and videos
3. Hands-on sessions using Python
4. Quiz/Puzzles
5. Seminars

MODULE-1

Introduction to Deep Learning

Introduction, Shallow Learning, Deep Learning, Why to use Deep Learning, How Deep Learning Works,Deep Learning Challenges,. How Learning Differs from Pure Optimization, Challenges in Neural Network Optimization.

Textbook 1: Ch 1.1 – 1.6, **Textbook 2:** 8.1,8.2

MODULE-2

Basics of Supervised Deep Learning

Introduction, Convolution Neural Network, Evolution of Convolution Neural Network, Architecture of CNN, Convolution Operation

Textbook 1: Ch 2.1 – 2.5

MODULE-3

Training Supervised Deep Learning Networks

Training Convolution Neural Networks, Gradient Descent-Based Optimization Techniques, Challenges in Training Deep Networks.

Supervised Deep Learning Architectures: LetNet-5,AlexNet

Text Book - 1 : Ch 3.2,3.4,3.5, Ch 4.2,4.3

MODULE-4

Recurrent and Recursive Neural Networks

Unfolding Computational Graphs, Recurrent Neural Network, Bidirectional RNNs, Deep Recurrent Networks, Recursive Neural Networks, The Long Short-Term Memory.Gated RNNs.

Text Book – 2: 10.1-10.3, 10.5, 10.6, 10.10

MODULE-5

Deep Reinforceme,nt Learning: Introduction, Stateless Algorithms: Multi-Armed Bandits, The Basic Framework of Reinforcement Learning, case studies.

Textbook – 3: Chapter 9: 9.1,9.2,9.3, 9.7

PRACTICAL COMPONENT OF IPCC*(May cover all / major modules)*

Sl.NO	Experiments
1	Design and implement a neural based network for generating word embedding for words in a document corpus
2	Write a program to demonstrate the working of a deep neural network for classification task.
3	Desing and implement a Convolutional Neural Network(CNN) for classification of image dataset
4	Build and demonstrate an autoencoder network using neural layers for data compression on image dataset.
5	Desing and implement a deep learning network for classification of textual documents.
6	Design and implement a deep learning network for forecasting time series data.
7	Write a program to enable pre-train models to classify a given image dataset
8	Simple Grid World Problem: Design a custom 2D grid world where the agent navigates from a start position to a goal, avoiding obstacles. Environment: Custom grid (easily implemented in Python)

Course outcomes (Course Skill Set):

At the end of the course, the student will be able to:

- CO1: Demonstrate the implementation ofdeep learning techniques
- CO2: Examine variousdeep learning techniquesfor solvingthe realworld problems
- CO3: Design and implementresearch-oriented scenario using deep learningtechniques in a team

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

CIE for the theory component of the IPCC (maximum marks 50)

- IPCC means practical portion integrated with the theory of the course.
- CIE marks for the theory component are **25 marks** and that for the practical component is **25 marks**.
- 25 marks for the theory component are split into **15 marks** for two Internal Assessment Tests (Two Tests, each of 15 Marks with 01-hour duration, are to be conducted) and **10 marks** for other assessment methods mentioned in 22OB4.2. The first test at the end of 40-50% coverage of the syllabus and the second test after covering 85-90% of the syllabus.
- Scaled-down marks of the sum of two tests and other assessment methods will be CIE marks for the theory component of IPCC (that is for **25 marks)**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the theory component of IPCC.

CIE for the practical component of the IPCC

- **15 marks** for the conduction of the experiment and preparation of laboratory record, and **10 marks** for the test to be conducted after the completion of all the laboratory sessions.

- On completion of every experiment/program in the laboratory, the students shall be evaluated including viva-voce and marks shall be awarded on the same day.
- The CIE marks awarded in the case of the Practical component shall be based on the continuous evaluation of the laboratory report. Each experiment report can be evaluated for 10 marks. Marks of all experiments' write-ups are added and scaled down to **15 marks**.
- The laboratory test **(duration 02/03 hours)** after completion of all the experiments shall be conducted for 50 marks and scaled down to **10 marks.**
- Scaled-down marks of write-up evaluations and tests added will be CIE marks for the laboratory component of IPCC for **25 marks**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the practical component of the IPCC.

SEE for IPCC

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours**)

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scoredby the student shall be proportionally scaled down to 50 Marks

The theory portion of the IPCC shall be for both CIE and SEE, whereas the practical portion will have a CIE component only. Questions mentioned in the SEE paper may include questions from the practical component.

Suggested Learning Resources:

Text Book:

1. M. Arif Wani Farooq Ahmad Bhat Saduf Afzal Asif Iqbal Khan, Advances in Deep Learning, Springer, 2020
2. Ian Goodfellow, Yoshua Bengio, Aaron Courville, "Deep Learning", MIT Press, 2016.
3. Charu C. Aggarwal, "Neural Networks and Deep Learning", Springer, 2018.

Reference book:

1. Bengio, Yoshua. "Learning deep architectures for AI." Foundations and trends in Machine Learning, 2009
2. N.D. Lewis, "Deep Learning Made Easy with R: A Gentle Introduction for Data Science", January 2016
3. Nikhil Buduma, "Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms", O'Reilly publications

Web links and Video Lectures (e-Resources):

https://cedar.buffalo.edu/~srihari/

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

i) Activity Description: Review paper on Deep learning applications.

ii) Implementation:

a. **Synopsis submission:** Students will submit the synopsis based on their interest.
b. **Presentation / Demo:** Students will present their review paper along with report in specified format**.**

Statistical Machine Learning For Data Science		Semester	7
Course Code	**BAD702**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	3:0:2:0	SEE Marks	50
Total Hours of Pedagogy	40 hours Theory + 8-10 Lab slots	Total Marks	100
Credits	04	Exam Hours	3
Examination nature (SEE)	Theory/practical		

Course objectives:

- Understand Exploratory Data Analysis
- Explain Data and Sampling Distributions
- To Analyse Statistical experiments and perform significance testing
- To demonstrate how to perform regression analysis on the data
- Explain Discriminant Analysis on the data.

Teaching-Learning Process (General Instructions)

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) need not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding.

MODULE-1

Exploratory Data Analysis: estimates of locations and variability, exploring data distributions, exploring binary and categorical data, exploring two or more variables.

Textbook : Chapter 1

MODULE-2

Data and Sampling Distributions: Random sampling and bias, selection bias, sampling distribution of statistic, bootstrap, confidence intervals, data distributions: normal, long tailed, student's-t, binomial, Chi-square, F distribution, Poisson and related distributions.

Textbook : Chapter 2

MODULE-3

Statistical Experiments and Significance Testing: A/B testing, hypothesis testing, resampling, statistical significance & p-values, t-tests, multiple testing, degrees of freedom.

Textbook : Chapter 3

MODULE-4

Multi-arm bandit algorithm, power and sample size, factor variables in regression, interpreting the regression equation, Regression diagnostics, Polynomial and Spline Regression.

Textbook : Chapter 3 & 4

MODULE-5

Discriminant Analysis: Covariance Matrix, Fisher's Linear discriminant, Generalized Linear Models, Interpreting the coefficients and odd ratios, Strategies for Imbalanced Data.

Textbook : Chapter 5

PRACTICAL COMPONENT OF IPCC

Sl.NO	Experiments
1	A dataset contains the prices of houses in a city. Find the 25th and 75th percentiles and calculate the interquartile range (IQR). How does the IQR help in understanding the price variability?
2	You are given a dataset with categorical variables about customer satisfaction levels (Low, Medium, High) and whether customers made repeat purchases (Yes/No). Create visualizations such as bar plots or stacked bar charts to explore the relationship between satisfaction level and repeat purchases. What can you infer from the data?
3	A dataset contains information about car models, including the engine size (in Liters), fuel efficiency (miles per gallon), and car price. Use a pair plot or correlation matrix to explore the relationships between these variables. Which variables seem to have the strongest relationships, and what might be the practical significance of these findings?
4	You want to estimate the mean salary of software engineers in a country. You take 10 different random samples, each containing 50 engineers, and calculate the sample mean for each. Plot the distribution of these sample means. How does the Central Limit Theorem explain the shape of this sampling distribution, even if the underlying salary distribution is skewed?
5	A researcher conducts an experiment with a sample of 20 participants to determine if a new drug affects heart rate. The sample has a mean heart rate increase of 8 beats per minute and a standard deviation of 2 beats per minute. Perform a hypothesis test using the t-distribution to determine if the mean heart rate increase is significantly different from zero at the 5% significance level.
6	A company is testing two versions of a webpage (A and B) to determine which version leads to more sales. Version A was shown to 1,000 users and resulted in 120 sales. Version B was shown to 1,200 users and resulted in 150 sales. Perform an A/B test to determine if there is a statistically significant difference in the conversion rates between the two versions. Use a 5% significance level.
7	You are comparing the average daily sales between two stores. Store A has a mean daily sales value of $1,000 with a standard deviation of $100 over 30 days, and Store B has a mean daily sales value of $950 with a standard deviation of $120 over 30 days. Conduct a two-sample t-test to determine if there is a significant difference between the average sales of the two stores at the 5% significance level.
8	A company collects data on employees' salaries and records their education level as a categorical variable with three levels: "High School", "Bachelor's", and "Master's". Fit a multiple linear regression model to predict salary using education level (as a factor variable) and years of experience. Interpret the coefficients for the education levels in the regression model.
9	You have data on housing prices and square footage and notice that the relationship between square footage and price is nonlinear. Fit a spline regression model to allow the relationship between square footage and price to change at 2,000 square feet. Explain how spline regression can capture different behaviours of the relationship before and after 2,000 square feet.
10	A hospital is using a Poisson regression model (a type of GLM) to predict the number of emergency room visits per week based on patient age and medical history. The model is given by: $\text{Log}(\lambda) = 2.5 - 0.03*\text{Age} + 0.5*\text{condition}$ where λ is the expected number of visits per week, **Age** is the patient's age, and **condition** is a binary variable (1 if the patient has a chronic condition, 0 otherwise). Interpret the coefficients of Age and condition. What is the expected number of visits per week for a 60-year-old patient with a chronic condition? How would the expected number of visits change if the patient did not have a chronic condition?
11	A bakery claims that its new cookie recipe is lower in calories compared to the old recipe, which had a mean calorie count of 200. You sample 40 new cookies and find a mean of 190 calories with a standard deviation of 15 calories. Perform a one-tailed t-test to determine if the new recipe has significantly fewer calories at a 5% significance level.

Course outcomes (Course Skill Set):

At the end of the course, the student will be able to:

- Analyse data sets using techniques to estimate variability, exploring distributions, and investigating relationships between variables.
- Apply random sampling, confidence intervals, and recognize various data distributions on datasets.
- Perform significance testing and identify statistical significance.
- Apply regression analysis for prediction, interpret regression equations, and assess regression diagnostics.
- Perform discriminant analysis on the varieties of datasets.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

CIE for the theory component of the IPCC (maximum marks 50)

- IPCC means practical portion integrated with the theory of the course.
- CIE marks for the theory component are **25 marks** and that for the practical component is **25 marks**.
- 25 marks for the theory component are split into **15 marks** for two Internal Assessment Tests (Two Tests, each of 15 Marks with 01-hour duration, are to be conducted) and **10 marks** for other assessment methods mentioned in 22OB4.2. The first test at the end of 40-50% coverage of the syllabus and the second test after covering 85-90% of the syllabus.
- Scaled-down marks of the sum of two tests and other assessment methods will be CIE marks for the theory component of IPCC (that is for **25 marks)**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the theory component of IPCC.

CIE for the practical component of the IPCC

- **15 marks** for the conduction of the experiment and preparation of laboratory record, and **10 marks** for the test to be conducted after the completion of all the laboratory sessions.
- On completion of every experiment/program in the laboratory, the students shall be evaluated including viva-voce and marks shall be awarded on the same day.
- The CIE marks awarded in the case of the Practical component shall be based on the continuous evaluation of the laboratory report. Each experiment report can be evaluated for 10 marks. Marks of all experiments' write-ups are added and scaled down to **15 marks**.
- The laboratory test **(duration 02/03 hours)** after completion of all the experiments shall be conducted for 50 marks and scaled down to **10 marks.**
- Scaled-down marks of write-up evaluations and tests added will be CIE marks for the laboratory component of IPCC for **25 marks**.
- The student has to secure 40% of 25 marks to qualify in the CIE of the practical component of the IPCC.

SEE for IPCC

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours**)

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored by the student shall be proportionally scaled down to 50 Marks

The theory portion of the IPCC shall be for both CIE and SEE, whereas the practical portion will have a CIE component only. Questions mentioned in the SEE paper may include questions from the practical component.

Suggested Learning Resources:

Books

1. Peter Bruce, Andrew Bruce and Peter Gadeck, "Practical Statistics for Data Scientists", 2nd edition, O'Reilly Publications, 2020.

Web links and Video Lectures (e-Resources):

Statistical learning for Reliability Analysis: https://nptel.ac.in/courses/106105239

Engineering Statistics: https://nptel.ac.in/courses/127101233

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

- Course (mini) project to demonstrate the concepts (10 marks)

	DATA SECURITY AND PRIVACY		Semester	7
	Course Code	**BAD703**	CIE Marks	50
	Teaching Hours/Week (L:T:P: S)	4:0:0:0	SEE Marks	50
	Total Hours of Pedagogy	50	Total Marks	100
	Credits	04	Exam Hours	3
	Examination type (SEE)	Theory		

Course objectives:

1. Understand the basics of, Security, its principle and Cryptography
2. To study varios symmetric and asymmetric cryptographic Algorithm
3. Apply the knowlwdge of Cryptography to various fields
4. Studty the key management system
5. Understand the necessity of data security

Teaching-Learning Process

These are sample Strategies, which teachers can use to accelerate the attainment of the various course outcomes.

1. Lecturer method (L) needs not to be only a traditional lecture method, but alternative effective teaching methods could be adopted to attain the outcomes.
2. Use of Video/Animation to explain functioning of various concepts.
3. Encourage collaborative (Group Learning) Learning in the class.
4. Ask at least three HOT (Higher order Thinking) questions in the class, which promotes critical thinking.
5. Adopt Problem Based Learning (PBL), which fosters students' Analytical skills, develop design thinking skills such as the ability to design, evaluate, generalize, and analyze information rather than simply recall it.
6. Introduce Topics in manifold representations.
7. Show the different ways to solve the same problem with different circuits/logic and encourage the students to come up with their own creative ways to solve them.
8. Discuss how every concept can be applied to the real world - and when that's possible, it helps improve the students' understanding
9. Use any of these methods: Chalk and board, Active Learning, Case Studies

Module-1 10 hours

A model for Network Security, Classical encryption techniques: Symmetric cipher model, Substitution ciphers-Caesar Cipher, Monoalphabetic Cipher, Playfair Cipher, Hill Cipher, Polyalphabetic Ciphers, One time pad, Steganography

Block Ciphers and Data Encryption Standards: Traditional Block Cipher structures, data Encryption Standard (DES), A DES Example, The strength of DES, Block cipher design principles

Text book 1:
Chapter 1: 1.8 Chapter 3: 3.1, 3.2, 3.5 Chapter 4: 4.1, 4.2, 4.3, 4.4, 4.5

Module-2 10 hours

	Pseudorandom number Generators: Linear Congruential Generators, Blum Blum Shub Generator Public key cryptography and RSA: Principles of public key cryptosystems-Public key cryptosystems, Applications for public key cryptosystems, Requirements for public key cryptography, Public key Cryptanalysis, The RSA algorithm: Description of the Algorithm, Computational aspects, The Security of RSA Diffie-Hellman key exchange: The Algorithm, Key exchange Protocols, Man-in-the-middle Attack, Elliptic Curve Cryptography: Analog of Diffie-Hellman key Exchange, Elliptic Curve Encryption/Decryption, Security of Elliptic Curve Cryptography Text book 1: Chapter 8: 8.2 Chapter 9: 9.1, 9.2 Chapter 10: 10.1, 10.4
	Module-3 10 hours
	Key management fundamentals, Key lengths and lifetimes, Key generation, Key establishment, Key storage, Key usage, Governing key management. **Public-Key Management**: Certification of public keys, The certificate lifecycle, Public-key management models, Alternative approaches. Text book 2: Chapter 10, Chapter 11
	Module-4 10 hours
	Web security consideration, Transport layer security. IP Security: IP Security overview, IP Security Policy, Encapsulating Security Payload, Combining security associations, Internet key exchange. Text book1: Chapter 17: 17.1, 17.2 Chapter 20: 20.1, 20.2, 20.3 20.4, 20.5
	Module-5 10 hours
	Data Security: Data hiding in Text-Basic features, Applications of data hiding, Watermarking, Intuitive Methods, Simple Digital methods, Data hiding in Text, Innocuous Text, Mimic Functions. Data hiding in Images: LSB encoding, BPCS Steganography, Lossless data hiding Textbook 3: Chapter 10: 10.1, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8 Chapter 11: 11.1, 11.2, 11.3

Course outcome

At the end of the course, the student will be able to :

CO1: Explain the basic concepts of Security and Cryptography

CO2: Analyze various Cryptographic Algorithm

CO**3:** Describe various key management scenarios.

CO4: Explain about IP security and Web security.

CO5: Apply the Data security concepts for Text and images.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

Continuous Internal Evaluation:

- For the Assignment component of the CIE, there are 25 marks and for the Internal Assessment Test component, there are 25 marks.
- The first test will be administered after 40-50% of the syllabus has been covered, and the second test will be administered after 85-90% of the syllabus has been covered
- Any two assignment methods mentioned in the 22OB2.4, if an assignment is project-based then only one assignment for the course shall be planned. The teacher should not conduct two assignments at the end of the semester if two assignments are planned.
- For the course, CIE marks will be based on a scaled-down sum of two tests and other methods of assessment.

Internal Assessment Test question paper is designed to attain the different levels of Bloom's taxonomy as per the outcome defined for the course.

Semester-End Examination:

Theory SEE will be conducted by University as per the scheduled timetable, with common question papers for the course (**duration 03 hours).**

1. The question paper will have ten questions. Each question is set for 20 marks.
2. There will be 2 questions from each module. Each of the two questions under a module (with a maximum of 3 sub-questions), **should have a mix of topics** under that module.
3. The students have to answer 5 full questions, selecting one full question from each module.
4. Marks scored shall be proportionally reduced to 50 marks

Suggested Learning Resources:

Books

Textbook

1. Cryptography and Network Security", William Stallings, Pearson Publication, Seventh Edition.

2. Everyday Cryptography: Fundamental Principles and Applications Keith M. Martin Oxford Scholarship Online: December 2013.

3. Data Privacy and Security, Salomon, David, Springer, 2003.

Reference Books:

1. Cryptography and Network Security, Behrouz A Forouzan, Dedeep Mukhopadhyay, TMH, 2nd edition, 2013
2. Information Security: Principles and practice, Mark Stamp, Wiley Inter Science, 2011

Activity Based Learning (Suggested Activities in Class)/ Practical Based learning

- Group assignments (TWO) to implement Cryptographic Algorithms (15+10 marks]