APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

(A State Government University)

B. Tech, 2024

Minor Degree in

Artificial Intelligence

Offered By : Computer Science and Engineering and Allied Branches

CURRICULUM

Minor (Artificial Intelligence)											
Sl. No:	Semester	Course Code	Course Title (Course Name)	Credit Structure			SS	Total Marks		Credits	Hrs./ Week
				L	T	P		CIA	ESE		
1	3	MNCST319	PYTHON FOR ARTIFICIAL INTELLIGENCE*	3	1	0	5	40	60	4	4
2	4	MNCST419	MATHEMATICS FOR MACHINE LEARNING*	3	1	0	5	40	60	4	4
3	5	MNCST519	ESSENTIALS OF MACHINE LEARNING*	3	1	0	5	40	60	4	4
4	6	MNCST619	DEEP LEARNING*	3	0	0	4.5	40	60	3	3
Total							**20/ 21**			**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.*

*&The courses offered in the third and fourth semesters can be structured as either theory-based courses **or** a combination of theory and lab-based courses.*

SYLLABUS

SEMESTER 3

SEMESTER 3

PYTHON FOR ARTIFICIAL INTELLIGENCE

Course Code	**MNCST319**	**CIE Marks**	40
Teaching Hours/Week (L:T:P)	3:1:0	ESE Marks	60
Credits	4	Exam Hours	2 Hrs. 30 Min.
Prerequisites (if any)	None	Course Type	**Theory**

Course Objectives:

- To provide learners an insight into Python programming in a scientific computation context and develop programming skills to solve engineering problems
- To give insights in developing AI & web based applications
- To familiarize Data Visualization tools & its application .

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Foundations of Coding** - Control statements - Selection structure (if-else), Iteration structure (for, while), Using built in functions and modules, Working with List, tuples. Sets, Dictionaries, Work with Functions -Arguments and return values, Lambda functions.	**11**
2	**Design with classes** - Objects and Classes, Methods, Instance Variables, Constructor, Accessors and Mutators. Structuring classes with Inheritance and Polymorphism. Abstract Classes. Exceptions - Handle a single exception, handle multiple exceptions.	**11**
3	**Packages for Numerical Operations**: NumPy - Basics, creating arrays, Arithmetic, Slicing, Matrix Operations, Random Numbers, Linear Algebra-Solving System of Equations **Plotting and visualization Packages**- Matplotlib, Basic plot, Ticks, Labels, and Legends, Ploting graph with Numpy and Matplotlib-line, scatter plot, bar chart, pie chart, sine and sigmoid functions etc, Basics of Seaborn package.	**11**

4	**Packages for data processing and prediction**: Pandas - Series and DataFrame, Reading Data from Files Using Pandas, Extracting Information from a DataFrame, Grouping and Aggregation, Data cleaning functions. **Implentation of Simple Regression Problem**-Reading and preprocessing the data from CSV file, Prediction using scikit-learn library.	**11**

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 =24marks)**	1. Each question carries 9 marks. 2. Two questions will be given from each module, out of which 1 question should be answered. 3. Each question can have a maximum of 3 sub divisions. **(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	Write, test and debug Python programs using conditional and iterative statements	**K3**
CO2	Develop programs by utilizing the modules Lists, Tuples, Sets and Dictionaries in Python and functions	**K3**
CO3	Implement Object Oriented programs with exception handling	**K3**
CO4	Write programs in Python to process data stored in files for predicting the results	**K3**

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

CO-PO Mapping Table:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	**3**		**3**		**3**							**3**
CO2	**3**		**3**		**3**							**3**
CO3	**3**	**3**	**3**		**3**							**3**
CO4	**3**	**3**	**3**		**3**							**3**

Text Books

Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Fundamentals of Python : First Programs	Kenneth A Lambert	Cengage Publishing	2/e, 2016
2	Python for Data Analysis,	Wes McKinney	Shroff / O'Reilly Publishers	2/e, 2017
3	Introduction to Python for Science and Engineering,	David J. Pine,	CRC Press	1/e, 2021
4	The Complete Reference Python	Martin C. Brown	Mc Graw Hill	1/e, 2018

Reference Books

Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Think Python: How to Think Like a Computer Scientist,	Allen B. Downey	Schroff	2/e, 2016
2	Python Programming, Shroff/Murach	Michael Urban and Joel Murac	Shroff/Murach, 2016	2016
3	Python Essential Reference.	David M.Baezly	Addison-Wesley Professional	4/e, 2009.
4	Python for Informatics: Exploring Information,	Charles Severance	Addison-Wesley Professional	2009

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1	The Joy of Computing using Python - Course By Prof. Sudarshan Iyengar, Prof. Yayati Gupta | IIT Ropar
2	Programming in Python - Course By Dr. Rizwan Rehman | Dibrugarh University

MODEL QUESTION PAPER

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

THIRD SEMESTER B. TECH MINOR DEGREE EXAMINATION, MONTH AND YEAR

Course Code: MNCST319

Course Name: PYTHON FOR ARTIFICIAL INTELLIGENCE

Max. Marks: 60		Duration: 2 hours 30 minutes

PART A

		Answer all questions. Each question carries 3 marks	CO	Mark s
1		Explain the syntax of **for loop** to display the even numbers within the range 2 and 100 in reverse order.	1	3
2		Write a Python script to find the factorial of a given number.	1	3
3		Discuss the use of __init__method in a class with suitable examples.	2	3
4		What is polymorphism? Give an example in the context of OOP in Python.	2	3
5		Describe random number generation using Python.	3	3
6		Write Python code to plot Sine wave.	3	3
7		Show creation the Series object with suitable example.	4	3
8		Discuss Simple Linear Regression model.	4	3

PART B

Answer any one full question from each module. Each question carries 9 marks

Module 1

9	a)	Discuss the relation between tuples, lists, and dictionaries in detail	1	5
	b)	Write a Python code to check whether a given year is a leap year or not [An year is a leap year if it's divisible by 4 but not divisible by 100 except for those divisible by 400]**.**	1	4
10	a)	Explain the use of Lambda function with suitable example.	1	4
	b)	Write a Python program to read a list of numbers and sort the list in a nondecreasing order without using any built in functions. Separate function should be written to sort the list wherein the name of the list is passed as the parameter.	1	5

Module 2				
11	a)	Explain inheritance in Python. Give examples for each type of inheritance	2	4
	b)	Write a Python class named Circle constructed by a radius and two methods which will compute the area and the perimeter of a given circle.	2	5
12	a)	How is exception handling accomplished in Python programs?	2	4
	b)	Write Python program to create a class called as Complex and implement add ( ) method to add two complex numbers. Display the result by overloading the + Operator.	2	5
Module 3				
13	a)	Describe the slicing operation of NumPy Array.	3	4
	b)	Write a Python program to add two matrices and also find the transpose of the resultant matrix.	3	5
14	a)	Discuss the syntax for plotting scatter plot and pie chart	3	4
	b)	Write python program to solve the following system of equations $x1 - 2x2 + 9x3 + 13x4 = 1$ $-5x1 + x2 + 6x3 - 7x4 = -3$ $4x1 + 8x2 - 4x3 - 2x4 = -2$ $8x1 + 5x2 - 7x3 + x4 = 5$	3	5
Module 4				
15	a)	Discuss the data pre-processing methods of Dataframe for replacing empty values and deleting the duplicate rows.	4	4
	b)	Given a file "auto.csv" of automobile data with the fields index, company, body-style, wheel- base, length, engine-type, num-of- cylinders, horsepower, average-mileage, and price, write Python codes using Pandas to • Load the data • Find the average mileage of all companies • Find the highest priced car of all companies.	4	5
16	a)	Explain the basic steps in creating a model.	4	4
	b)	Write a Python code to predict value of y for x=6, using simple linear regression model trained using the following dataset. x: 1.0, 2.0, 3.0, 4.0, 5.0 y: 1.00, 2.00, 1.30, 3.75, 2.25	4	5
*****				

SEMESTER 4

SEMESTER 4

MATHEMATICS FOR MACHINE LEARNING

Course Code	MNCST419	CIE Marks	40
Teaching Hours/Week (L:T:P)	3:1:0	ESE Marks	60
Credits	4	Exam Hours	2 Hrs. 30 Min.
Prerequisites (if any)	Higher secondary mathematics	Course Type	Theory

Course Objectives:

1. To lay down the basic concepts and techniques of linear algebra, calculus and basic probability theory needed for subsequent study.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**LINEAR ALGEBRA :** Systems of Linear Equations – Matrices, Solving Systems of Linear Equations. Vector Spaces - Linear Independence, Basis and Rank, Linear Mappings, Norms, - Inner Products - Lengths and Distances - Angles and Orthogonality - Orthonormal Basis - Orthogonal Complement - Orthogonal Projections. Matrix Decompositions – Determinant and Trace, Eigenvalues and Eigenvectors, Cholesky Decomposition, Eigen decomposition and Diagonalization, Singular Value Decomposition, Matrix Approximation.	11
2	**VECTOR CALCULUS :** Differentiation of Univariate Functions - Partial Differentiation and Gradients, Gradients of Vector Valued Functions, Gradients of Matrices, Useful Identities for Computing Gradients. Back propagation and Automatic Differentiation – Higher Order Derivatives- Linearization and Multivariate Taylor Series.	11
3	**Probability and Distributions:** Construction of a Probability Space - Discrete and Continuous Probabilities, Sum Rule, Product Rule, and Bayes' Theorem. Summary Statistics and Independence – Important Probability distributions - Conjugacy and the Exponential Family - Change of Variables/Inverse Transform	11
4	**Optimization:** Optimization Using Gradient Descent - Gradient Descent With Momentum, Stochastic Gradient Descent. Constrained Optimization and Lagrange Multipliers – Convex Optimization - Linear Programming - Quadratic Programming.	11

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 =24marks)**	• 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 sub divisions. **(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	Make use of the concepts, rules and results about linear equations, matrix algebra, vector spaces, eigenvalues & eigenvectors and orthogonality & diagonalization to solve computational problems.	**K3**
CO2	Perform calculus operations on functions of several variables and matrices, including partial derivatives and gradients	**K3**
CO3	Utilize the concepts, rules and results about probability, random variables, additive & multiplicative rules, conditional probability, probability distributions and Bayes' theorem to find solutions of computational problems	**K3**
CO4	Train Machine Learning Models using unconstrained and constrained optimization methods	**K3**

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

CO-PO Mapping Table:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	**3**	**3**										**3**
CO2	**3**	**3**										**3**
CO3	**3**	**3**										**3**
CO4	**3**	**3**										**3**

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Mathematics for Machine Learning.	Deisenroth, M.P., Faisal, A.A., Ong, C. S.	Cambridge University Press, UK	2020

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Linear Algebra and Its Applications	Gilbert Strang	Cengage Learning, United States	4/e,2006
2	Convex Optimization	Stephen Boyd and Lieven Vandenberghe	Cambridge University Press	2018

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1	**https://nptel.ac.in/courses/111107137**

MODEL QUESTION PAPER

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B. TECH MINOR DEGREE EXAMINATION, MONTH AND YEAR

Course Code: MNCST419

Course Name: MATHEMATICAL FOUNDATIONS FOR MACHINE LEARNING

Max. Marks: 60		Duration: 2 Hours 30 Minutes

PART A

		Answer all questions. Each question carries 3 marks	CO	Marks
1		Check the linear independence of the vectors (2,-1,3), (1, 1,-2), (3, 3, -8)	1	(3)
2		Is $(Z,.)$ a group? Justify your answer	1	(3)
3		For a scalar function $f(x,y,z) = x^2 + 3y^2 + z^2$, find the gradient and its magnitude at the point (1,2,-1).	2	(3)
4		Find the Taylor polynomials $T_n, n = 0,1,2$ of $f(x) = cosx - sinx$	2	(3)
5		Let A and B be events such that P(A)=0.35, P(B)=0.40, and $P(A \cup B) =$ 0.5, find $P(A \vee B)$.	3	(3)
6		What is an exponential family? Why are exponential families useful?	3	(3)
7		Explain the principle of the gradient descent algorithm.	4	(3)
8		Briefly explain the difference between (batch) gradient descent and stochastic gradient descent. Give an example of when you might prefer one over the other	4	(3)

PART B

Answer any one full question from each module. Each question carries 9 marks

Module 1

9		Use Gram-Schmidt process to find an orthogonal basis for the column space of the matrix $\begin{pmatrix} 2 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$	1	9
10		Find the singular value decomposition (SVD) of the matrix $\begin{pmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$	1	9

Module 2

11	a)	Find the critical points of $f(x) = x^2 - 3xy + 5x - 2y + 6y^2 + 8$	2	4
	b)	Find the linear approximation to the function $f(x,y) = 2 - sin\,(-x - 3y)$ at the point $(0, \pi)$, and then use it to estimate $f(0.001, \pi)$.	2	5
12	a)	Find the maximum and minimum values of the function $f(x) = 4x + 4y - x^2 - y^2$ subject to the condition $x^2 + y^2 \leq 2$	2	4

<table>
<tr><td></td><td>b)</td><td>Let $f(x,y)=\begin{cases}\frac{x^2y}{x^2+y^2},(x,y)\neq 0\\ 0,(x,y)=0\end{cases}$
(a) Find the partial derivatives of $f(x,y)$
(b) Show that $f(x,y)$ is not differentiable at (0,0)</td><td>2</td><td>5</td></tr>
<tr><td colspan="5">Module 3</td></tr>
<tr><td rowspan="2">13</td><td>a)</td><td>Given a continuous random variable x, with cumulative distribution function $f_x(x)$, show that the random variable $y=f_x(x)$ is uniformly distributed.</td><td>3</td><td>4</td></tr>
<tr><td>b)</td><td>There are two bags. The first bag contains four mangos and two apples; the second bag contains four mangos and four apples. We also have a biased coin, which shows “heads” with probability 0.6 and “tails” with probability 0.4. If the coin shows “heads”, we pick a fruit at random from bag 1; otherwise we pick a fruit at random from bag 2. Your friend flips the coin (you cannot see the result), picks a fruit at random from the corresponding bag, and presents you a mango. What is the probability that the mango was picked from bag 2?</td><td>3</td><td>5</td></tr>
<tr><td rowspan="2">14</td><td>a)</td><td>Roll a fair dice twice. Let the random variable X be the product of the outcomes of the two rolls.
(a) What is the probability mass function of X? .
(b) What are the expected value and the standard deviation of X?</td><td>3</td><td>5</td></tr>
<tr><td>b)</td><td>A random variable X has the following probability distribution.
<table><tr><td>x</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>$P(X=x)$</td><td>0.2</td><td>a</td><td>b</td><td>0.24</td><td>0.16</td></tr></table>(i) Given E(X)=0.25, find the values of a,b
(ii) Find $P(X>2)$</td><td>3</td><td>4</td></tr>
<tr><td colspan="5">Module 4</td></tr>
<tr><td>15</td><td></td><td>Find the maximum value of $f(x,y,z)=xy+z$, given that $g(x,y,z)=x+y+z=3, x\geq 0, y\geq 0, z\geq 0$</td><td>4</td><td>9</td></tr>
<tr><td rowspan="2">16</td><td>a)</td><td>Is the function $f(x,y)=2x^2+y^2+6xy-x+3y-7$ convex, concave or neither? Justify.</td><td>4</td><td>4</td></tr>
<tr><td>b)</td><td>Consider the univariate function $f(x)=x^3+6x^2-3x-5$. Find its stationary points and indicate whether they are maximum, minimum, or saddle points.</td><td>4</td><td>5</td></tr>
<tr><td colspan="5">*****</td></tr>
</table>

SEMESTER 5

SEMESTER 5

ESSENTIALS OF MACHINE LEARNING

Course Code	MNCST519	CIE Marks	40
Teaching Hours/Week (L:T:P)	3:1:0	ESE Marks	60
Credits	4	Exam Hours	2 Hrs. 30 Min.
Prerequisites (if any)	None	Course Type	Theory

Course Objectives:

1. To equip the learners to understand the basic and advanced concepts and algorithms in machine learning.
2. To enable the learners to use standard and most popular supervised and unsupervised learning algorithms.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Introduction to ML -** Machine Learning vs. Traditional Programming, Machine learning paradigms - supervised, semi-supervised, unsupervised, reinforcement learning. **Basics of parameter estimation** - maximum likelihood estimation (MLE) and maximum aposteriori estimation (MAP), Bayesian formulation. **Supervised Learning -** Feature Representation and Problem Formulation, Role of loss functions and optimization **Regression** - Linear regression with one variable, Linear regression with multiple variables - solution using gradient descent algorithm and matrix method.	**10**
2	**Classification** - Naive Bayes, KNN **Generalisation and Overfitting** - Idea of overfitting, LASSO and RIDGE regularization, Idea of Training, Testing, Validation **Evaluation Measures –** Classification - Precision, Recall, Accuracy, F-Measure, Receiver Operating Characteristic Curve(ROC), Area Under Curve (AUC). Regression - Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R Squared/Coefficient of Determination.	**8**

3	**Neural Networks (NN) -** Perceptron, Neural Network - Multilayer feed-forward network, Activation functions (Sigmoid, ReLU, Tanh), Back propagation algorithm, Issues in Back Propagation Learning (Vanishing and Exploding Gradient Problems, Local Minima and Saddle Points, Overfitting, Slow Convergence, Sensitivity to the selection of the Initial weights, Computational Cost), Use of Adaptive Learning Rate Methods (Back Propagation Algorithm with Momentum, Adaptive Gradient Descent Algorithm (ADAGRAD), Root Mean Square Propagation (RMSProp), Adaptive Momentum Estimation (ADAM)) **Decision Trees –** Information Gain, Gain Ratio, ID3 algorithm	**13**
4	**Clustering** - Similarity measures, Hierarchical Clustering - Agglomerative Clustering, partitional clustering, K-means clustering **Dimensionality reduction:-** Principal Component Analysis, Multidimensional scaling **Ensemble methods:-** Bagging, Boosting **Resampling methods:-** Bootstrapping, Cross Validation. Practical aspects - Bias-Variance trade-off	**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 =24marks)**	• 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 sub divisions. **(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	Illustrate Machine Learning concepts and basic parameter estimation methods	K2
CO2	Demonstrate supervised learning concepts (regression, classification)	K3
CO3	Illustrate the concepts of Multilayer neural network and Decision trees	K3
CO4	Describe unsupervised learning concepts and dimensionality reduction techniques	K3
CO5	Use appropriate performance measures to evaluate machine learning models	K3

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

CO-PO Mapping Table:

	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**
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**	**3**							**2**
CO5	**3**	**3**	**3**	**3**	**3**							**2**

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	4/e, 2020
2	Data Mining and Analysis: Fundamental Concepts and Algorithms	Mohammed J. Zaki,, Wagner Meira	Cambridge University Press	1/e, 2016

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Machine Learning	Tom Mitchell	McGraw-Hill	1/e, 1997
2	Neural Networks for Pattern Recognition	Christopher Bishop	Oxford University Press	1/e, 1995
3	Machine Learning: A Probabilistic Perspective	Kevin P Murphy	MIT Press	1/e, 2012
4	The Elements of Statistical Learning	Trevor Hastie, Robert Tibshirani, Jerome Friedman	Springer	2/e, 2007
5	Foundations of Machine Learning	Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar	MIT Press	1/e, 2012

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1	https://youtu.be/fC7V8QsPBec?si=8kqBn-7x1RG5V1J
2	https://youtu.be/g__LURKuIj4?si=Xj10NPfMfpQSOhVx
3	https://youtu.be/yG1nETGyW2E?si=ySlxpeWuFAUQBf7-
4	https://youtu.be/zop2zuwF_bc?si=W7TpSHLdi4rykva4

MODEL QUESTION PAPER

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FIFTH SEMESTER B. TECH MINOR DEGREE EXAMINATION, MONTH AND YEAR

Course Code: MNCST519

Course Name: ESSENTIALS OF MACHINE LEARNING

Max. Marks: 60 | Duration: 2 hours 30 minutes

PART A

		Answer all questions. Each question carries 3 marks	CO	Marks
1		Explain how machine learning differs from traditional programming.	1	(3)
2		Classify the following learning paradigms into supervised, unsupervised, semi-supervised, and reinforcement learning with a brief definition for each.	1	(3)
3		Explain the difference between training, testing, and validation in the context of a machine learning model.	2	(3)
4		Summarize the common evaluation metrics used to measure the performance of a classification model.	2	(3)
5		Explain the role of activation functions in a neural network. Give examples of any two commonly used activation functions.	3	(3)
6		Summarize the issues that arise in training neural networks using backpropagation.	3	(3)
7		Explain the concept of similarity measures in clustering. Mention any two commonly used similarity/distance measures.	4	(3)
8		Summarize the difference between **bagging** and **boosting** in ensemble learning.	4	(3)

PART B

Answer any one full question from each module. Each question carries 9 marks

Module 1

9	a)	Explain how gradient descent helps in optimizing the loss function in supervised learning.	1	3
	b)	Apply the gradient descent algorithm for one iteration to minimize the cost function in a linear regression model given the data points: (1, 2), (2, 3), (3, 5), (4, 7). Assume initial parameters $\theta_0 = 0, \theta_1 = 1$ and learning rate $\alpha = 0.01$.	1	6
10	a)	Explain how the Bayesian formulation aids in improving parameter estimation when dealing with uncertain or limited data.	1	3
	b)	Analyze the difference between Maximum Likelihood Estimation and Maximum A Posteriori Estimation in terms of: (i) Objective functions (ii) Use of prior information. Use a Gaussian distribution or a coin toss example to support your explanation.	1	6

<table>
<tr><th colspan="5">Module 2</th></tr>
<tr><td>11</td><td>a)</td><td>Explain how Naive Bayes handles feature independence in classification. Why is this assumption useful in practice?</td><td>2</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Apply the K-Nearest Neighbors classification technique to classify a test point $x = 6$using the following training data:
<table>
<tr><th>x</th><th>Class</th></tr>
<tr><td>1</td><td>A</td></tr>
<tr><td>3</td><td>A</td></tr>
<tr><td>7</td><td>B</td></tr>
<tr><td>8</td><td>B</td></tr>
</table></td><td>2</td><td>6</td></tr>
<tr><td>12</td><td>a)</td><td>Compare and contrast the following regression evaluation metrics with appropriate interpretation:
(i) Mean Absolute Error
(ii) Root Mean Squared Error
(iii) R^2- Value</td><td>2</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Analyze the problem of overfitting in machine learning models.
(i) What are the symptoms and consequences of overfitting?
(ii) How do LASSO and RIDGE regularization address this issue?</td><td>2</td><td>6</td></tr>
<tr><th colspan="5">Module 3</th></tr>
<tr><td>13</td><td>a)</td><td>Explain how ReLU differs from Sigmoid regarding output behaviour and suitability in deep networks.</td><td>3</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Apply the Perceptron learning rule for one iteration to learn the OR function using initial weights $w_0 = 0, w_1 = 0, w_2 = 0$, learning rate $\eta = 1$, and input vector (1,1) with target output 1. Show all intermediate steps.</td><td>3</td><td>6</td></tr>
<tr><td>14</td><td>a)</td><td>Compare Information Gain and Gain Ratio as splitting criteria in decision trees. Why is Gain Ratio considered an improvement over Information Gain?</td><td>3</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Analyze the challenges faced in backpropagation learning, especially focusing on the vanishing gradient and exploding gradient problems. Also, describe how ADAM and RMSProp address these challenges.</td><td>3</td><td>6</td></tr>
<tr><th colspan="5">Module 4</th></tr>
<tr><td>15</td><td>a)</td><td>Explain how Principal Component Analysis helps in dimensionality reduction.</td><td>4</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Apply K-means clustering to the following data points: (1,1), (2,1), (4,3), (5,4) with $k = 2$. Use Manhattan distance and initial centroids as (1,1) and (5,4). Perform one iteration and show cluster assignments.</td><td>4</td><td>6</td></tr>
<tr><td>16</td><td>a)</td><td>Compare agglomerative hierarchical clustering and partitional clustering regarding algorithmic structure and output type.</td><td>4</td><td>3</td></tr>
<tr><td></td><td>b)</td><td>Analyze how bootstrapping and cross-validation help in model evaluation. Also discuss their impact on the bias-variance trade-off.</td><td>4</td><td>6</td></tr>
<tr><td colspan="5"></td></tr>
</table>

SEMESTER 6

SEMESTER 6

DEEP LEARNING

Course Code	MNCST619	CIE Marks	40
Teaching Hours/Week (L:T:P)	3:0:0	ESE Marks	60
Credits	3	Exam Hours	2 Hrs. 30 Min.
Prerequisites (if any)	None	Course Type	Theory

Course Objectives:

1. To give the learner an understanding about the foundations of Deep Learning architecture and applications
2. To equip the learner with the necessary skills to set-up neural network architecture and use it for real time problem solution.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	The Basic Architecture of Neural Networks - Single Computational Layer: The Perceptron, Example: Learning OR, AND. Multilayer Neural Networks. Activation functions – Sign, Sigmoid, Tanh, ReLU, leaky ReLU, Hard Tanh, Softmax. Example: Learning XOR, Loss function. Training a Neural Network with Backpropagation. Practical issues in neural network training. Overfitting, Underfitting, Capacity, Hyper parameters and Validation sets, Estimators -Bias and Variance.	9
2	Deep feed forward network, setup and initialization issues, Vanishing and exploding gradient problems, Optimization techniques - Gradient Descent (GD), Stochastic GD, GD with momentum, GD with Nesterov momentum, AdaGrad, RMSProp, Adam. Regularization Techniques - L1 and L2 regularization, Early stopping, Dataset augmentation, Parameter tying and sharing, Ensemble methods, Dropout.	9
3	Convolutional Neural Networks –Architecture, Motivation, padding. Efficient convolution algorithms, Applications of Convolutional Networks, Pre-trained convolutional Architectures : AlexNet, ResNet. Recurrent neural networks – Computational graphs. RNN design. Encoder – decoder sequence to sequence architectures. Language modeling example of RNN. Gated RNNs LSTM and GRU.	9
4	Autoencoders, Variational AutoEncoder, Applications of Autoencoders. Deep generative models - Boltzmann machines, Restricted Boltzmann Machines, Deep Belief Networks, Deep Boltzmann Machines, Generative Adversarial Networks. Transfer Learning and Domain	9

	Adaptation. Computer Vision - Speech Recognition	

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 =24marks)**	• 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 sub divisions. **(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	Construct fundamental neural network architectures and algorithms, including Multilayer Perceptron and Back-propagation	K3
CO2	Apply standard regularization and optimization techniques for the effective training of deep neural networks.	K3
CO3	Build convolutional Neural Network (CNN) and Recurrent Neural Networks (RNNs) models for different use cases.	K3
CO4	Apply Autoencoders, deep generative models, Generative Adversarial Networks (GANs)and transfer learning to solve complex problems in computer vision and speech recognition.	K3

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

CO-PO Mapping Table:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	**3**	**3**	**2**		**3**							**3**
CO2	**3**	**3**	**2**		**3**							**3**
CO3	**3**	**3**	**2**		**3**							**3**
CO4	**3**	**3**	**2**		**3**							**3**

Text Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Neural Networks and Deep Learning	Charu C. Aggarwal	Springer	2018
2	The Science of Deep Learning	Iddo Drori	Cambridge University Press	2023
3	Deep Learning	Ian Goodfellow, Yoshua Bengio, Aaron Courville	MIT Press	2016

Reference Books				
Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Learning Deep Architectures for AI	Yoshua Bengio	Now Publishers	1/e, 2009
2	Deep Learning: A Practitioner's Approach	Josh Patterson, Adam Gibson	O'Reilly	1/e, 2017
3	Applied Deep Learning. A Case-based Approach to Understanding Deep Neural Networks	Umberto Michelucci	Apress	1/e, 2018
4	Deep Learning with Keras	Antonio Gulli, Sujit Pal	Packt	1/e, 2017
5	Deep Learning with Python	Francois Chollet	Manning	1/e. 2017
6	Deep Learning	M Gopal	Pearson	1/e, 2022

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1	https://nptel.ac.in/courses/106105215 (Week 4)
2	https://nptel.ac.in/courses/106105215 (Week 5)
3	https://nptel.ac.in/courses/106105215 (Week 8)
4	https://nptel.ac.in/courses/106105215 (Week 10,11 and 12)

MODEL QUESTION PAPER

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

SIXTH SEMESTER B. TECH MINOR DEGREE EXAMINATION, MONTH AND YEAR

Course Code: MNCST619

Course Name: DEEP LEARNING

Max. Marks: 60		Duration: 2 hours 30 minutes

PART A

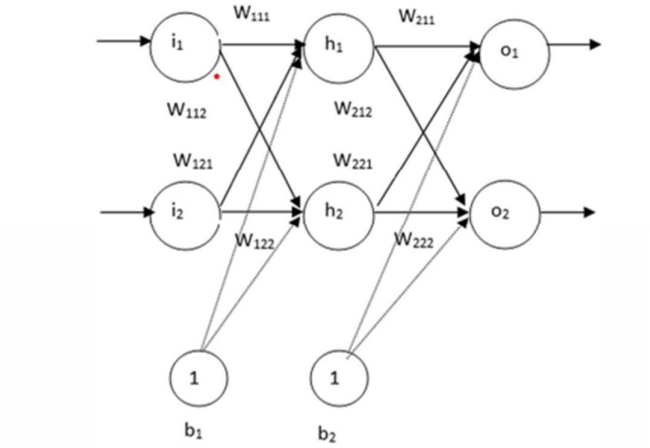
		Answer all questions. Each question carries 3 marks	CO	Marks
1		Sketch the typical learning curves for the training and validation sets, for a setting where overfitting occurs at some point. Assume that the training set and the validation set are of the same size.	1	(3)
2		Illustrate the limitation of a single layer perceptron with an example.	1	(3)
3		Initializing the weights of a neural network with very small or large random numbers is not advisable. Justify.	2	(3)
4		A given function is of the form J(Θ)=Θ^2 -Θ+2.What is the weight update rule for the gradient descent optimization at step t+1? Consider α=0.01 to be the learning rate	2	(3)
5		Determine the shape of output matrix of an image of size 19 × 19 that uses a padding size 2, stride size 2, and a 5 × 5 filter.	3	(3)
6		If we have a recurrent neural network (RNN), we can view it as a different type of network by "unrolling it through time". Justify.	3	(3)
7		How does the variational auto-encoder architecture allow it to generate new data points, compared to auto-encoder, which cannot generate new data points?	4	(3)
8		Is an autoencoder for supervised learning or for unsupervised learning? Explain.	4	(3)

PART B

Answer any one full question from each module. Each question carries 9 marks

Module 1

9	a)	A 2 × 2 image is represented by the following pixel value matrix. $\begin{bmatrix} 5 & 4 \\ 2 & 7 \end{bmatrix}$ This image is given to a 4-layer neural network (one input layer, two hidden layers and one output layer). Draw schematic diagram of the network. Assuming all inter-connection weights having values 1, bias	1	3

		having value 0, the hidden layers having 3 neurons each, and a simple activation function of the form $\frac{1}{1+x}$ being used, **compute** output for one round of forward propagation.		
	b)	Implement AND function using perceptron networks for Bipolar inputs and Bipolar targets.	1	6
10	a)	Consider a fully connected neural network with one hidden layer and two output nodes. The network takes a two-dimensional input vector x = [x1, x2] as input, and activation function used in both the hidden and output layers is the sigmoid activation function. The weight and biases of the network are initialized randomly as follows: W1 = [[w111 = 0.2, w112 = -0.4], [w121 = 0.1, w122 = -0.1]], b1 = [-0.3, 0.4] W2 = [w211 = 0.3, w212 = 0.1], [w221 = - 0.2, w222 = 0.2]], b2 = [-0.2, 0.1]  Consider a training example with input x = [1,2], and target output y =[0.9,0.1]. Compute the gradient with respect to one of the weights w222 of the network using back propagation, assuming a mean squared error loss function. (Take η =0.5)	1	9
Module 2				
11	a)	**Derive** the gradient descent training rule assuming for the target function $o_d = w_0 + w_1x_1 + ... + w_nx_n$. Define explicitly the squared cost/error function E, assuming that a set of training examples D is provided, where each training example d ε D is associated with the target output t_d.	2	4
	b)	Training error of the deep learning model trained for the classification problem was found to be very low but generalization error was high. Identify the problem and suggest techniques to reduce this generalization error.	2	5
12	a)	Differentiate gradient descent with and without momentum. Give equations for weight updation in GD with and without momentum. Illustrate plateaus, saddle points and slowly varying gradients.	2	6

	b)	Describe the effect in bias and variance when a neural network is modified with more number of hidden units followed with dropout regularization.	2	3
Module 3				
13	a)	In Convolutional Neural Networks, there is no need to perform feature extraction. Justify with an example.	3	6
	b)	Weight sharing allows CNNs to deal with image data without using too many parameters. Does weight sharing increase the bias or the variance of a model?	3	3
14	a)	The vanishing gradient problem is more pronounced in RNN than in traditional neural networks. Discuss a solution for the problem.	3	5
	b)	Show the steps involved in an Long short Term Memory (LSTM) to predict stock prices.	3	4
Module 4				
15	a)	Generative Adversarial Networks (GANs) include a generator and a discriminator. Sketch a basic GAN using those elements, a source of real images, and a source of randomness	4	5
	b)	The word "adversarial" in the acronym for GANs suggests a two-player game. What are the two players, and what are their respective goals?	4	4
16	a)	Anu has trained a classifier for flowers and performs very well. Bob decided to build his own model using transfer learning. What are the additional hyperparameters (due to the transfer learning) Bob will need to tune? Justify.	4	5
	b)	Is it possible to use Generative Adversarial Network (GAN) for data augmentation? Justify.	4	4
*****				