HINDUSTAN
INSTITUTE OF TECHNOLOGY & SCIENCE
(DEEMED TO BE UNIVERSITY)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CURRICULUM AND SYLLABUS

Under CBCS

(Applicable for Students admitted from Academic Year 2018-19)

B. Tech. Computer Science Engineering

With Specialization in Artificial Intelligence and Machine Learning

(In Collaboration with IBM)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING SCIENCES

HINDUSTAN INSTITUTE OF TECHNOLOGY & SCIENCE
VISION AND MISSION

MOTTO

"To Make Every Man A Success And No Man A Failure"

VISION

To be an International Institute of Excellence, providing a conducive environment for education with a strong emphasis on innovation, quality, research and strategic partnership blended with values and commitment to society.

MISSION

- To create an ecosystem for learning and world class research.
- To nurture a sense of creativity and innovation.
- To instill highest ethical standards and values with a sense of professionalism.
- To take up activities for the development of Society.
- To develop national and international collaboration and strategic partnership with industry and institutes of excellence.
- To enable graduates to become future leaders and innovators.

VALUE STATEMENT

- Integrity, Innovation, Internationalization

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
VISION AND MISSION

VISION

To excel in Computer Science and Engineering education, research and project management by empowering the students with strong conceptual knowledge.

MISSION

M1: To educate the students with basic foundation blocks of core and allied disciplines of Computer Science and Engineering.

M2: To provide practical skills in the advancements of the Computer Science and Engineering field required for the growing dynamic IT and ITES industries.

M3: To sculpt strong personal, technical, research, entrepreneurial, and leadership skills.

M4: To inculcate knowledge in lifelong learning, professional ethics and contribution to the society.

B. Tech. Computer Science and Engineering

PROGRAMME EDUCATIONAL OBJECTIVES (PEO)

The Program Educational Objectives (**PEOs**) of the **Computer Science and Engineering** are listed below:

The graduate after 3-5 years of programme completion will

PEO1: Excel in his/her professional career and/or pursue higher education including research by applying the knowledge of Computer Science and Engineering.

PEO2: Demonstrate the technical skills to analyze and design appropriate solutions for problems with social consciousness and ethical values.

PEO3: Adapt themselves to organizational needs by understanding the dynamically changing technologies.

PROGRAM OUTCOMES (ALIGNED WITH GRADUATE ATTRIBUTES) (PO)

(To be achieved by the student after every semester/year/and at the time of graduation)

At the end of this program, graduates will be able to

PO1: **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO2: **Problem analysis:** Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3: **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4: **Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5: **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8: **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9: **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: **Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12: **Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAM SPECIFIC OUTCOMES (PSO)

On completion of the B.Tech. Computer Science & Engineering degree the graduates will be able to

PSO1: Apply mathematical, conceptual knowledge of computing and analytical skills to solve complex problems.

PSO2: Design and develop computer systems based on the domains of Cyber Physical Systems, Algorithm Design Techniques and Enterprise systems security.

PSO3: Do innovative system design with analytical knowledge by developing modern tools and techniques.

HINDUSTAN INSTITUTE OF TECHNOLOGY & SCIENCE (DEEMED TO BE UNIVERSITY)

SCHOOL OF COMPUTING SCIENCES

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)									
(172 CREDIT STRUCTURE)									
SEMESTER – I									
SL. NO	**COURSE CATEGORY**	**COURSE CODE**	**NAME OF THE COURSE**	**L**	**T**	**P**	**C**	**S**	**TCH**
1	HS/ES	ELA4101/ MEA4101	Professional English and Soft Skills /Engineering Graphics and Computer Aided Design	1	1	2	3	1	4
2	BS	MAA4101	Matrices and Calculus	3	0	2	4	0	5
3	BS	PHA4102/CYA4101	Engineering Physics/Engineering Materials	3	0	0	3	1	3
4	PC	CSA4101	Problem Solving Using C	2	0	2	3	1	4
5	ES	EEB4101 /CSB4101	Introduction to Digital Systems / Engineering and Design	3	0	0	3	1	3
6	ES	GEA4131	Engineering Immersion Lab	0	0	2	0.5	2	2
7	BS	PHA4131/CYA4131	Engineering Physics Lab/ Materials Chemistry Lab	0	0	2	1	0	2
Total				12	1	10	**17.5**	6	23

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)									
(172 CREDIT STRUCTURE)									
SEMESTER – II									
SL. NO	**COURSE CATEGORY**	**COURSE CODE**	**NAME OF THE COURSE**	**L**	**T**	**P**	**C**	**S**	**TCH**
1	BS	MAA4117	Analytical Mathematics	3	0	2	4	1	5
2	BS	PHA4102/CYA4101	Engineering Physics/ Engineering Materials	3	0	0	3	1	3

3	HS/ES	ELA4101/ MEA4101	Professional English and Soft Skills /Engineering Graphics and Computer Aided Design	1	1	2	3	1	4
4	ES	EEB4101 / CSB4101	Introduction to Digital Systems / Engineering and Design	2	0	2	3	1	4
5	ES	GEA4102	Sustainable Engineering Systems	2	0	0	2	1	3
6	PC	CSB4117	Data Structures using C	3	0	0	3	1	3
7	PC	CSB4118	Object Oriented Programming using C++	3	0	2	4	1	5
8	PC	CSB4146	Data Structures Lab	0	0	3	1	0	3
9	ES	GEA4131	Engineering Immersion Lab	0	0	2	0.5	2	2
10	BS	PHA4131/CYA4131	Engineering Physics Lab/ Materials Chemistry Lab	0	0	2	1	0	2
11	**PC**	**IBA4101**	**Introduction to Artificial Intelligence and Machine Learning**	**3**	**0**	**0**	**3**	**0**	**3**
Total				20	1	15	**27.5**	9	37

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)									
(172 CREDIT STRUCTURE)									
SEMESTER – III									
SL. NO	**COURSE CATEGORY**	**COURSE CODE**	**NAME OF THE COURSE**	**L**	**T**	**P**	**C**	**S**	**TCH**
1	BS	MAA4201	Partial Differential Equations and Transforms	3	0	2	4	0	5
2	PC	CSB4201	Design and Analysis of Algorithms	2	1	2	4	1	5
3	PC	CSB4202	Database Management Systems	3	0	0	3	1	3
4	PC	CSB4203	Java Programming	3	0	2	4	0	5
5	**DE**	**IBA4201**	**Applied statistics**	**3**	**0**	**2**	**4**	**1**	**5**
6	NE	CSD42**	Non-Department Elective- I	2	0	0	2	0	2
7	**PC**	**IBA4202**	**Introduction to Python Programming**	**3**	**0**	**2**	**4**	**0**	**5**
8	PC	CSB4232	Database Management Systems Lab	0	0	3	1	0	3
Total				19	1	13	**26**	3	33

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)									
(172 CREDIT STRUCTURE)									
SEMESTER – IV									
SL. NO	**COURSE CATEGORY**	**COURSE CODE**	**NAME OF THE COURSE**	**L**	**T**	**P**	**C**	**S**	**TCH**
1	BS	MAA4219	Discrete Mathematics	3	1	0	4	0	4
2	PC	CSB4216	Computer Organization and Architecture	3	0	0	3	1	3
3	PC	CSB4217	Computer Networks	3	0	0	3	1	3
4	PC	CSB4218	Operating Systems	3	0	0	3	1	3
5	**DE**	**IBA4302**	**Data Mining and Predictive Modelling**	**3**	**0**	**2**	**4**	**0**	**5**
6	NE	CSD42**	Non-Department Elective–II	2	0	0	2	0	2
7	PC	CSB4241	Networking Lab	0	0	3	1	0	3
8	PC	CSB4242	Operating Systems Lab	0	0	3	1	0	3
9	PC	CSB4243	Design Project-I	0	0	2	1	0	2
10	-	-	Internship	0	0	0	1	0	0
Total				17	1	10	**23**	3	28

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)									
(172 CREDIT STRUCTURE)									
SEMESTER – V									
SL. NO	**COURSE CATEGORY**	**COURSE CODE**	**NAME OF THE COURSE**	**L**	**T**	**P**	**C**	**S**	**TCH**
1	BS	MAA4302	Probability and Statistics	3	0	2	4	0	5
2	PC	CSB4301	Web Technology	2	0	2	3	1	4
3	PC	CSB4302	Theory of Computation	3	1	0	4	1	4
4	**PC**	**IBA4401**	**Application of Machine Learning in Industries**	**2**	**0**	**2**	**3**	**0**	**3**
5	HS	GEA4216	Professional Ethics and Life Skills	2	0	0	2	1	2
6	**DE**	**IBA4216**	**Machine Learning**	**3**	**0**	**2**	**4**	**0**	**5**

7	NE	CSD43**	Non-Department Elective–III	2	0	0	2	0	2
8	PC	CSB4331	Skill Development in Programming	0	0	2	1	0	2
9	PC	CSB4332	Design Project with IoT	0	0	3	1	0	3
Total				17	1	13	**24**	3	30

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)

(172 CREDIT STRUCTURE)

SEMESTER – VI

SL. NO	COURSE CATEGORY	COURSE CODE	NAME OF THE COURSE	L	T	P	C	S	TCH
1	PC	CSB4316	Principles of Compiler Design	3	1	0	4	1	4
2	**PC**	**IBA4316**	**Deep Learning**	**3**	**0**	**2**	**4**	**0**	**5**
3	PC	CSB4318	Data Warehousing and Data Mining	3	0	0	3	1	3
4	PC	CSB4319	Modern Software Engineering	3	0	2	4	1	5
5	HS	GEA4304	Business Economics	2	0	0	2	1	2
6	NE	CSD43**	Non-Department Elective–IV	2	0	0	2	0	2
7	PC	CSB4341	Compiler Design lab	0	0	3	1	0	3
8	PC	CSB4342	Design Project-II	0	0	2	1	0	2
9	-	-	Internship	0	0	0	1	0	0
Total				16	1	9	**22**	4	26

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)

(172 CREDIT STRUCTURE)

SEMESTER – VII

SL. NO	COURSE CATEGORY	COURSE CODE	NAME OF THE COURSE	L	T	P	C	S	TCH
1	PC	CSB4401	Software Project Management	3	0	0	3	1	3
2	PC	CSB4402	Big Data and Analytics	3	0	2	4	1	5

3	PC	CSB4403	Applied Cryptography and Network Security	3	1	0	4	1	4
4	PC	CSB4404	Programming Paradigms	3	0	0	3	1	3
5	**PC**	**IBA4301**	**Computational Linguistics and Natural Language Processing**	**3**	**0**	**2**	**4**	**0**	**5**
6	DE	CSD44**	Non-Department Elective-V	2	0	0	2	0	2
7	PC	CSB4431	Cloud Deployment Lab	2	0	2	3	0	4
8	**PC**	**IBA4431**	**Mini Project**	**0**	**0**	**2**	**1**	**0**	**2**
Total				19	1	8	**24**	4	28

B.TECH. CSE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (IN COLLABORATION WITH IBM)

(172 CREDIT STRUCTURE)

SEMESTER – VIII

SL. NO	COURSE CATEGORY	COURSE CODE	NAME OF THE COURSE	L	T	P	C	S	TCH
1	PC	CSB4441	Project & Viva – voce	0	0	16	8	0	16
Total				0	0	16	**8**	0	16
Total							**172**		

SYLLABUS

SEMESTER – II

COURSE TITLE	INTRODUCTION TO ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING			CREDITS	3
COURSE CODE	IBA4101	COURSE CATEGORY	PC	L-T-P-C	3-0-0-3
CIA	50%			ESE	50%
LEARNING LEVEL	BTL-3				

CO	COURSE OUTCOMES	PO
Upon completion of this course, the students will be able to		
1	Understand the concept of Artificial Intelligence	1,2
2	Familiarize with Knowledge based AI systems and approaches	1,2,5,12
3	Apply the aspect of Probabilistic approach to AI	1,2,5,12
4	Identify the Neural Networks and NLP in designing AI models	1,2,3,5,12
5	Recognize the concepts of Machine Learning and its deterministic tools	1,2,5,12

MODULE 1: INTRODUCTION TO ARTIFICIAL INTELLIGENCE (6)

History of artificial intelligence, The birth of artificial intelligence, AI Winters, Todays' AI, Historical milestones in the development of AI, Great contributors, People who have influenced AI , Differences between strong AI and weak AI, Artificial Intelligence definitions, Emergence of AI – Technological advances, Machine Learning ---> Deep Learning --->AI, Functions of AI, Characteristics of artificial intelligence, Applications of AI, AI in health care, Industry 4.0, AI in manufacturing, AI in education sector, AI in business, AI in Finance Sector, AI in Law, AI in society, Cognitive science and AI, Cognition and process of Cognition, Disciplines in Cognitive science, Multidisciplinary subject, Linguistics, Artificial intelligence as Cognitive science, Methods in Cognitive science, Watson.

MODULE 2: LOGICAL APPROACH TO AI AND KNOWLEDGE-BASED SYSTEM (6)

Introduction to knowledge representation systems, Knowledge representation using logic, Propositional logic, Semantics of propositional logic, Properties of propositional logic statements, Tautologies and logical implication, Resolution, Conjunctive normal form, Resolution is valid, Resolution algorithm, Knowledgebase systems, Structure of a knowledge based system, Recap of artificial intelligence, Components of expert systems, Expert systems development, Wumpus world, Logic, A simple knowledge base, Exploring the Wumpus world, Semantic net, Inference in semantic networks, Semantic networks: Types and components, Types of relationships in semantic network, Frames, Frames: Some examples, Non-monotonic logic, Circumscription, Default logic.

MODULE 3: PROBABILISTIC APPROACH TO AI (9)

Probability, Basic concepts, Probability of an event, Example on Sample Space, counting rules, Event relations, Conditional Probabilities, Defining Independence, The Law of Total Probability, Bayes' Rule,

Examples. Random Variables, Discrete Random Variable, Probability Distributions, Probability Mass Function, Probability Density Function, Expectations of Random Variables, Medians of Random Variables, The variance of a Random Variable, Chebyshev's Inequality, Quantiles of Random Variables, Jointly Distributed Random Variables, Marginal Probability Distributions, Independence and Covariance, Bayesian Networks, Merits of Bayesian Networks, Construction of a Bayesian Network, Representation in Bayesian Networks, Benefits of Bayesian Networks, Why learn Bayesian networks? Constructing Bayesian networks, Example from medical diagnostics, Software for Bayesian networks, Gaussian Bayesian Networks, Linear Gaussian BN to joint Gaussian, Theorem: From Gaussian to Bayesian networks, Noisy OR-Gate model, Promedas: A clinical diagnostic decision support system, Organization of PROMEDAS development.

MODULE 4: EVOLUTIONARY INTELLIGENCE AND NEURAL NETWORKS, NATURAL LANGUAGE UNDERSTANDING (12)

Using LSTMs to synthesize text, Neural Style transfer and applications, Image synthesis with variational auto encoders, Generative Adversarial Networks: 1 What does a GAN look like? 2 Generator, 3 Discriminator, 4. Generator vs Discriminator, 5 Training GANs.

MIMO Deep Learning models, Layers graphs: acyclic and directional, can layers share weights? Hyper parameter tuning, Bag of tricks: ensemble of models.

Introduction, Artificial Neural Network, Appropriate problems for neural network learning, Characteristics of the problems, Basic understanding of neural networks, A single neuron, Activation Functions, Architectures of neural networks, Feedforward neural network, Single-Layer feedforward architecture, Multiple-Layer feedforward architecture, Types of feedforward networks, Multi-layer perceptron, Training MLP: The back-propagation algorithm, Step 1: Forward propagation, Step 2: Back propagation and weight updation, Process of learning in neural network , Recurrent or feedback architecture, Mesh Architectures, GRADIENT-DESCENT(training examples, η), Stochastic GRADIENTDESCENT(training examples, η), Multilayer networks and Backpropagation algorithm, The Backpropagation algorithm, Natural language processing, Classical NLP, Feed-forward networks, Recurrent neural networks and recursive networks, Features for NLP problems, Framenet Vs. Wordnet, Features for text, Features for word relations, NGRAM features, Some terminologies.

MODULE 5: INTRODUCTION TO MACHINE LEARNING AND LEARNING DETERMINISTIC MODELS (12)

Motivation for Machine Learning, Applications, Machine Learning, Learning associations, Classification, Regression, The Origin of machine learning, Uses and abuses of machine learning, Success cases, How do machines learn, Abstraction and knowledge representation, Generalization, Factors to be considered, Assessing the success of learning, Metrics for evaluation of classification method, Steps to apply machine learning to data, Machine learning process, Input data and ML algorithm, Classification of machine learning algorithms, General ML architecture, Group of algorithms, Reinforcement learning, Supervised learning, Unsupervised learning, Semi-Supervised learning, Algorithms, Ensemble learning, Matching data to an appropriate algorithm.

Supervised Learning, Regression, Linear regression, Multiple linear regression, A multiple regression analysis, The analysis of variance for multiple regression, Examples for multiple regression, Overfitting,

Detecting overfit models: Cross validation, Cross validation: The ideal procedure, Parameter estimation, Logistic regression, Decision trees: Background, Decision trees, Decision trees for credit card promotion, An algorithm for building decision trees, Attribute selection measure: Information gain, Entropy, Decision Tree: Weekend example, Occam's Razor, Converting a tree to rules, Unsupervised learning, Semi-Supervised learning, Clustering, K – means clustering, Automated discovery, Reinforcement learning, Multi-Armed Bandit algorithms, Influence diagrams, Risk modelling, Sensitivity analysis, Casual learning.

TEXT BOOK	
1.	Introduction to Artificial Intelligence and Machine Learning (IBM ICE Publications).

SEMESTER- III

COURSE TITLE	**APPLIED STATISTICS**			**CREDITS**	**4**
COURSE CODE	**IBA4201**	**COURSE CATEGORY**	**PC**	**L-T-P-C**	**3-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-4**				

CO	**COURSE OUTCOMES**	**PO**
Upon completion of this course, the students will be able to		
1	Appreciate the Statistical Analysis concepts with their relationships and process	1,2
2	Familiarize with describing data, transforming and summarizing.	1,2,5,12
3	Design and develop R programs for testing hypothesis with real time applications.	1,2,3,5,12
4	Apply the examining relationships to find the correlation and regression.	1,2,5,12
5	Perform advanced techniques with applications of decision tress, neural networks	1,2,3,5,12

Prerequisites : NIL

MODULE 1: INTRODUCTION TO STATISTICAL ANALYSIS **(12)**
Introduction, Meaning of Statistics, The Scientific Method, Basic Steps of the Research Process, Experimental Data and Survey Data, Populations and Samples, Census and Sampling Method, Parameter and Statistic, Independent and Dependent Variables, Examining Relationships, Introduction to SPSS Statistics.
MODULE 2: DESCRIBING DATA **(12)**
Introduction, Types of Data, Data Transformation, Summarizing Data: Graphical Methods, Summarizing Data: Measures of Central Tendency, Summarizing Data: Measures of Dispersion, Levels of Measurement, Random Variables and Probability Distributions, Discrete and Continuous Random Variable, Making Inferences about Populations from samples, Estimator and Estimate, Confidence Interval for Population Mean (Large Sample).
MODULE 3: TESTING HYPOTHESIS **(12)**

Introduction, Null and Alternative Hypothesis, Type I and Type II Error, The Procedure of Hypothesis Testing, Hypothesis Testing of a Population Mean: Large Sample, Hypothesis Testing of a Population Mean: Small Sample, Hypothesis Test of a Proportion (One Sample), Hypothesis Test of Population Variance, Hypothesis Test of Population Mean: Two Independent Samples(), Hypothesis Test of Population Mean: Dependent Samples (Paired Samples), Hypothesis Test about Two Population Proportion, Hypothesis Test about Two Population Variances, Analysis of Variance (ANOVA), Nonparametric Test, Sign Test for Paired Data, Wilcoxon Matched Pairs Signed Ranks Test (for n>10 pairs), Mann-Whitney U Test, Kruskal-Wallis Tests (H Test).
MODULE 4 : EXAMINING RELATIONSHIPS (12)
Introduction, Types of Correlation, Karl Pearson Coefficient Correlation, Spearman's Rank Order Correlation, Partial Correlation, Residuals and Plots, Simple Linear Regression, Multiple Regression Model, Repeated Measures, Non-linear Regression, Polynomial Regression Models, Weighted Least Squares, Two Stage Least Squares 1, Structural Equation Modeling.
MODULE 5: ADVANCED TECHNIQUES (12)
Identifying Groups: Classification, Probit Analysis, Discriminant Function Analysis, Proportional Odds Models, Decision Trees, Neural Networks, Cluster Analysis, Factor Analysis, Multidimensional Scaling.
TEXT BOOK
1. Applied Statistics (IBM ICE Publications)

COURSE TITLE	**INTRODUCTION TO PYTHON PROGRAMMING**			**CREDITS**	**4**
COURSE CODE	**IBA4202**	**COURSE CATEGORY**	**PC**	**L-T-P-C**	**3-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-4**				

CO	**COURSE OUTCOMES**	**PO**
Upon completion of this course, the students will be able to		
1	Appreciate the Python programming concept with simple examples	1,2
2	Familiarize with Python data types and regular expressions aspects.	1,2,5,12
3	Design and develop python programs, with Python installation, with key focus on files and exception handling.	1,2,3,5,12
4	Appreciation of nuances such as standard libraries of Python	1,2,5,12
5	Perform testing with applications of Python in IoT	1,2,3,5,12

Prerequisites : NIL
MODULE 1: INTRODUCTION TO PYTHON (12)
History of Python, What's Python buzz? Why to Learn Python? Features of Python, Python interpreter, Growth, Strength of Python, Python use cases, Flavors of Python, Python 2 vs Python 3, When would one choose 2.7? Introduction to Python Versions, Python Installation, Testing and Validating Installation, PIP3 LIST.

Practical Component:

Hardware requirement:

- i5 Processor, 8GB RAM, Internet Connection

Software Environment:

- IDE recommended PYCHARM (Recommended), JUPYTER, VISUAL STUDIO

(i) Python Program to Reverse a Given Number and compare it to the original number
(ii) Python Program to Read a Number n and Print the Natural Numbers
(iii) Python Program to Determine all Pythagorean Triplets in the Range
(iv) Python Program to Check if a Number is an Armstrong Number

MODULE 2: DATA TYPES & REGULAR EXPRESSION (12)

Simple Python Program, Basic Data Types Statement, Basic Data types identifiers, Basic Data types, Integer Data Type, Float and Complex Data Type, Mathematical Function, String Data Types, String Manipulation Function, Sample String Python Code, Data types List, List Basic Operation, List Sorting, List – Count & Append, Dictionary, Python-Specialized data type, Basic Data types Collections, Files Directories & Flow control, Expression, Unicode, Example of RE, Control Statements, Range Function.

Practical Component:

(i) Python Program to Remove Vowels and Punctuation from String
(ii) Python program to count the number of strings where the string length is 2 or more and the first and last character are same from a given list of strings
(iii) Python program to read a list, sorted in increasing order by the last element in each tuple from a given list of non-empty tuples
(iv) Python program to generate and print a list of first and last 5 elements where the values are square of numbers between a range
(v) Python program to merge two dictionaries

MODULE 3: PYTHON VERSIONS OF PYTHON INSTALLATION (12)

File & Directory manipulation, File functions, File object attributes, close ( ) method, Opening a binary file, File Attributes, read(read_fixed_size) readline( )tell(), Two Input modes to read data from Keyboard, Parallel Processing os.fork() and os.exec (), Function arguments pass by value and reference, Default arguments, Try Except Code, Types of Exceptions, Sample class Example, Instantiating & Inheriting classes, Serializing Python Objects.

Practical Component:

(i) Python program to find the second lowest grade of any student(s) from the given names and grades of each student using lists and lambda
(ii) Python program to convert a roman numeral to an integer
(iii) Python program to find validity of a string of parentheses, '(', ')', '{', '}', '[' and ']
(iv) Python program to calculate student Grades with Multiple Inheritance
(v) Python program to create an Address Book with Multiple Inheritance

MODULE 4 : STANDARD LIBRARIES (12)

Python Standard libraries, Creating modules, Adding variables in the Module, Order of search for the modules in the code, Standard Libraries (SYS Module), Data Compression and Archiving, Creating zip files using the ZipFile Object, Tar file tools, Tar file Exceptions, Gzip, Bzip2, Python Web Development, Request module: Access web page, Example of a code that cannot connect to the remote, Request module: passing the payload, Request module: To view the status code, Response Module, http Module: Overview, http Module : http client, Django - Web Development framework, Django Interface, Django - Creating an app, Flask: Web Development Framework, Flask, Flask Code, Flask -Templates, Flask - Template Engine : Jinga, Flask – Jinga Template creation, Flask – Jinga Template, Network programming, Extending and Embedding Python using C, Custom Function, Mapping Functions, Initialization Function, Building and Installing Extensions.

Practical Component:

(i) Python program to create an application to implement a Calculator
(ii) Python program to implement Euclidean Algorithm to compute the greatest common divisor (gcd)
(iii) Python program that reads a given expression and evaluates it
(iv) Python program to group a sequence of key-value pairs into a dictionary of lists
(v) Python program to convert an array to an array of machine values and vice versa
(vi) Python program to Create widget to draw a line, two buttons, insert a text and delete section of the text
(vii) Python program to create a calculator using GUI – TK
(viii) Python program to display details about the operating system, working directory, files and directories in the current directory, lists the files and all directories, scan and classify them as directories and files
(ix) Python program to get information about the file pertaining to the file mode and to get time values with components using local time and gmtime

MODULE 5: PYTHON IN IoT (12)

Python in IoT, Python on Rasbian image, RPIGPIO module, Paho -MQTT, Python code using MQTT libraries, Client Code: Publish and Subscribe, Some of the APIs explained in larger scope, Output, Numpy, Working with Numpy, Constructing Numpy arrays, Printing arrays, Arithmetic Operations on ma trix's, numpyzeros( ), Slicing the arrays, Random number generation, Load data from a text file, Mysqldb, Installation of Python Interface module, Creating a database, Tested Database, Python : PyMySQL module, Exception handling in pymysql, OpenCV, Open CV - Installation, Sample code – Open CV, Matplotlib, Matplotlib-Installation & Sample code, Sample code – Bar Chart, Pandas : Data Processing Module, Pandas-Installation, Pandas-Data structure, Pandas-Series data, Pandas – Series data Sample Code, Pandas-Data Frames, Pandas Basic functionality, Example code of a 1-dimensional list.

Practical Component:

(i) Python program to connect to Google using socket programming
(ii) Python program to perform Array operations using Numpy package

(iii)	Python program to display multiple types of charts using Matplotlib package
(iv)	Python program to perform File Operation on Excel DataSet

TEXT BOOK	
1.	Programming with Python (IBM ICE Publications 2018 Edition).

SEMESTER-IV

COURSE TITLE	**DATA MINING AND PREDICTIVE MODELING**			**CREDITS**	**4**
COURSE CODE	**IBA4302**	**COURSE CATEGORY**	**DE**	**L-T-P-C**	**3-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-3**				

CO	**COURSE OUTCOMES**	**PO**
Upon completion of this course, the students will be able to		
1	Recognize the process of formulating business objectives, data selection/collection, preparation and process to successfully design, build, evaluate and implement predictive models for a various business application.	1,2
2	Compare and contrast the underlying predictive modeling techniques.	1,2,5,12
3	Select appropriate predictive modeling approaches to identify particular cases	1,2,3,5,12
4	Appreciate the nuances of Support Vector Machines and clustering techniques.	1,2,5,12
5	Apply predictive modeling approaches using a suitable package such as SPSS Modeler	1,2,3,5,12

MODULE 1: DATA UNDERSTANDING & PREPARATION **(9)**
Identifying business objectives, translating business objectives to data mining goals, reading data from various sources – Database/ Excel/ Text/others, data visualization – tabular & graphic, distributions and summary statistics, field reordering, Reclassify data.
MODULE 2: DATA TRANSFORMATIONS **(9)**
Data quality issues, Data Audit, anomalies, relationships among variables, Extent of Missing Data, Segmentation, Outlier detection, Variable transformations, Variable derivation, Variable selection, Automated Data Preparation, combining data files, data restructuring, Aggregation, Duplicates removal, Sampling cases, Data Caching, Partitioning data, Missing Value replacement.
MODULE 3: MODELING TECHNIQUES – I **(9)**
Partitioning The Data - Training, Validation & Testing, Model selection, Model development techniques - Linear regression, Logistic regression, Discriminant analysis, Bayesian networks, Neural networks, Rule Induction.
MODULE 4: MODELING TECHNIQUES – II **(9)**
Support vector machines, Cox regression, Time series analysis, Decision trees, Clustering, Association Rules, Sequence Detection, Which Technique to use when.

MODULE 5: MODEL EVALUATION & DEPLOYMENT (9)
Model Validation, Determining Model Accuracy, Rule Induction Using CHAID, Automating Models for Categorical Targets, Automating Models for Continuous Targets, Comparing and Combining Models, Evaluation Charts for Model Comparison, Using Propensity Scores, Meta-Level Modeling, Error Modeling, Deploying Model, Exporting Model Results, Assessing Model Performance, Updating A Model. **Practical Component:** (i) Introduction to the Case Study (ii) Data Understanding and Preparation (iii) Partitioning Data
TEXT BOOK
1. Data Mining & Predictive Modeling (IBM ICE Publications).

SEMESTER-V

COURSE TITLE	**APPLICATION OF MACHINE LEARNING IN INDUSTRIES**			**CREDITS**	**3**
COURSE CODE	**IBA4401**	**COURSE CATEGORY**	**DE**	**L-T-P-C**	**3-0-0- 3**
CIA	**50%**			**ESE**	**50%**
LEARNING LEVEL	**BTL-3**				

CO	**COURSE OUTCOMES-**	**PO**
Upon completion of this course, the students will be able to		
1	Understand the concept of Machine Learning.	1,2
2	Familiarize with applications of Machine Learning in Banking sectors.	1,2,5,12
3	Appreciate the various applications in Communication and Education sectors.	1,2,5,12
4	Identify the applications in Health care and Government sectors .	1,2,5,12
5	Recognize the applications in Manufacturing, Transportation and Logistics sectors.	1,2,5,12

Prerequisites : Machine Learning

MODULE 1: MACHINES LEARNING IN BANKING AND SECURITIES (6)
Why machine learning in banking sector, Use of AI in banking and finance, Fraud detection, Tough competition in banking industry, Risk modelling and investment banks, Customer data management, Decreased customer experience and loyalty, Personalized marketing, Role of machine learning: Challenges of banking sector and securities, Widely used machine learning algorithms in banking and security, Fraud prevention and detection systems, Rule based and machine learning based approach in fraud detection, Anomaly detection: Ways to expose suspicious transactions in banks, Advanced fraud detection systems, Risk management systems, Case study: Application of machine learning for

financial risk management, Credit risk analysis using machine learning classifier, Investment prediction systems, Portfolio management systems, Objectives of portfolio management, Algorithmic trading, Deep learning for customer services, Chatbot: Deep learning approach, AI powered marketing systems, Deep learning in cyber security, Types of cyber-attacks in banks, Deep learning methods used in cyber security, Deep learning v/s restricted Boltzmann machines, Convolution Neural Networks (CNNs), Recurrent neural networks, Machine learning techniques: Loan underwriting & sentiment/news analysis, Sentiment or news analysis, Current challenges and opportunities: Banking and security domain.

MODULE 2: MACHINE LEARNING IN COMMUNICATION, MEDIA, HEALTHCARE AND LIFE SCIENCE (9)

Machine learning in communication, media and entertainment, Usage of machine learning in media and entertainment industry, Machine learning techniques for customer sentiment analysis, World embedding's, Sentiment analysis with long short term memory networks, Real-time analytics in communication, media and entertainment industries, Real time analytics and social media, Deep learning for social media analytics, Recommendations engines, Collaborative filtering, Memory based collaborative filtering, Model based collaborative filtering, Content based filtering, Hybrid recommendation systems, Summary of recommendation systems, Deep learning techniques on recommender systems.

Applications of machine learning in health and life sciences, The most important applications of machine learning in healthcare, Role of machine learning in drug discovery, Medical image analysis, Why deep learning for medical image analysis, Neural network and deep learning architecture, Comparisons between architecture of different types of deep learning models, Machine learning in genetics and genomics, Genomics and AI background, Two category of genomics, How to use deep learning effectively, Interpreting deep learning models, Predictive medicine: Prognosis and diagnostics accuracy, Predictive medicine: Examples, ML applications in breast cancer diagnosis and prognosis.

MODULE 3: MACHINE LEARNING IN EDUCATION, MANUFACTURING AND PETROLEUM INDUSTRIES (9)

Advantages of machine learning in education, learning analytics, Academic analytics, Action research, Educational data mining, Recommender system, Personalized adaptive learning, Learning analytics process, Data environment: What? Stakeholders: Who? Methods: How? Case study: Sentimental analysis for student's feedback using ML, Recommender systems in education, Domain model, Learner model, Students classification algorithm, Recommendation model, Case study: Application of ML in predicting students' performance, Proposed methodology, Data description, Sample data sets, Visualization, Selection of machine learning technique.

Introduction, Applications of machine learning in manufacturing industry, Deep learning for smart manufacturing, Machine learning for quality control in manufacturing, Case study, Construction of CNN, Experimental results, Efficiency of CNN for defect detection, Comparative experiments, Machine learning for fault assessment, Time frequency methods, Spectrograms: Short-Time Fourier Transform (STFT), Scalograms: Wavelet transform, Hilbert-Huang transform, Proposed CNN architecture for fault classification based on vibration signals, Case study 1, Machinery failure prevention technology.

MODULE 4 : MACHINE LEARNING IN GOVERNMENT ADMINISTRATION AND INSURANCE INDUSTRIES (9)

Introduction, Risk and compliance, Type of government problems appropriate for AI applications, AI for citizen services use cases, Answering questions, Routing requests, Translation, Drafting documents, Chat bots for communication between citizen and government, Media richness theory, Chatbots in the public sector, Case study, Data management services, Knowledge processing services, Application services, An application scenario, Classifications of citizen complaints using ML, Case study, Step 1: Document collection, Step 2: Prepossessing, Step 3: Feature extraction, Term frequency- Inverse document frequency, Step 4: Feature selection, Step 5: Classification, How to implement, Result.

Importance of machine learning in insurance, Potential use cases of machine leaning in insurance industry, Case study on insurance climb analysis using machine learning algorithms, Case study on using machine learning for insurance pricing optimization, Personalized marketing in insurance industry, Predictive model for insurance underwriting, Case study: Risk prediction in life insurance industry.

MODULE 5: MACHINE LEARNING IN RETAIL AND SUPPLY CHAIN, TRANSPORTATION AND LOGISTICS, ENERGY AND UTILITIES (12)

Introduction, Inventory management, Few use case examples, Benefits of predictive analytics to retailers, Robots-seeing to customer satisfaction, IoT: Prevention first, Predictive analytics: Weathering demand, Analysing buying patterns, Analysing traffic patterns, Assortment planning, Eliminate guess work, Feed the right stores, Get better information, Assortment planning to drive supply chain, Retail analytics, Domestic forecasting, Case study: Forecasting seasonal footwear demand using ML, Demand forecasting methods, Predictor variables in demand forecasting, Traditional techniques v/s machine learning techniques, Methodology, Machine learning techniques used, List of attributes from the aggregated data by month at the style level, Feature selection and engineering, List of attributes for feature selection, Dataset partitioning, Model building, Three step model, K-means clustering, Three steps followed in classification, Three sub-steps in prediction, Performance measurement, Results, Three step model,
Machine learning for supply chain management, Recommended architecture for machine learning models, Machine learning models use case.

Introduction, Applications of ML and artificial intelligence in transportation, Applications of machine learning in transport, Incident detection, Predictive models, Application of AI in aviation and public transportation, Aviation, Shared mobility, Buses, Intelligent urban mobility, Autonomous vehicles, Autonomous transportation, Artificial intelligence use cases in logistics, Back office AI, Cognitive customs, Predictive logistics, Predictive risk management, Seeing thinking and speaking logistics operations, ML powered customer experience, Limitations of AI techniques in transportation, Computation complexity of AI algorithms.

Introduction, Smart grid, Smart grid technologies, Key characteristics of smart grid, Machine learning applications in smart grid, Machine learning techniques for renewable energy generation, Forecasting

renewable energy generation, Wind power generation, Solar energy generation, Hydro power generation, Determining plant location, size and configuration, Managing renewable energy-integrated smart grid, Machine learning applications in wind energy forecasting, Case study: Wind power forecasting based on daily wind speed data, Wind energy output calculations based on hourly wind speed, Machine learning techniques used, LASSO regression, KNN regression, xGBoost regression, Random forest regression, Support vector regression, Wind power forecasting method using machine learning algorithm, About data set, Case studies : Case 1: Wind power forecasting based on daily mean wind speed and standard deviation, Forecasting accuracy of algorithms, Case 2: Wind power forecasting based on only daily mean wind speed, Case 3: Wind power forecasting for a different region.
TEXT BOOK
1. Application of machine learning in industries (IBM ICE Publications).

COURSE TITLE	**MACHINE LEARNING**			**CREDITS**	**4**
COURSE CODE	**IBA4216**	**COURSE CATEGORY**	**DE**	**L-T-P-S**	**3-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-4**				

CO	**COURSE OUTCOMES**	**PO**
Upon completion of this course, the students will be able to		
1	Understand the concept of Machine Learning.	1,2
2	Familiarize with Simple Linear Regression and Logistic Regression.	1,2,5,12
3	Appreciate the various nuances of Multiple Regressions and Model Building.	1,2,5,12
4	Identify and apply the Classification algorithms.	1,2,5,12
5	Apply the Clustering algorithms for developing applications.	1,2,3,5,12

Prerequisites : Data Mining

MODULE 1: INTRODUCTION TO MACHINE LEARNING (12)

The Origins of Machine Learning, Uses and Abuses of Machine Learning, How do Machines Learn? - Abstraction and Knowledge Representation, Generalization, Assessing the Success of Learning, Steps to Apply Machine Learning to Data, Choosing a Machine Learning Algorithm - Thinking about the Input Data, Thinking about Types of Machine Learning Algorithms, Matching Data to an Appropriate Algorithm.

Practical Component:

Hardware requirement:

- i5 Processor, 8GB RAM, Internet Connection

Software Environment:

- IDE recommended PYCHARM (Recommended), JUPYTER, VISUAL STUDIO

i. Introduction to pandas

ii. Introduction to NumPy
iii. Wine Quality Prediction
iv. Housing Price Prediction
v. Air Quality Prediction
vi. Bank Marketing
vii. Liver Disease Prediction
viii. Heart Disease Prediction
ix. Credit Default Prediction
x. Car Price Prediction
xi. Media Content Problem
xii. Online Retail Case Study
xiii. Airline Passengers Prediction
xiv. Energy Efficiency Analysis
xv. Stock Price Prediction
xvi. Car Evaluation
xvii. Movie Sentiment Analysis

MODULE 2: SIMPLE LINEAR REGRESSION (12)

Introduction to Simple Linear Regression, Simple Linear Regression Model Building, Estimation of Parameters Using Ordinary Least Squares, Interpretation of Simple Linear Regression Coefficients, Validation of Simple Linear Regression Model, Coefficient of Determination (R-squared) and Adjusted R-Squared, Spurious Regression, Hypothesis Test for Regression Coefficients (t-Test), Test for Overall Model: Analysis of Variance (F-Test), Residual Analysis.

Practical Component:

(i) Program to demonstrate Simple Linear Regression
(ii) Program to demonstrate Logistic Regression using SCIKIT learn
(iii) Program to demonstrate Logistic Regression

MODULE 3: MULTIPLE REGRESSION AND MODEL BUILDING (12)

Introduction, Ordinary Least Squares Estimation for Multiple Linear Regression, Multiple Linear Regression Model Building, Partial Correlation and Regression Model Building, Interpretation of Multiple Linear Regression Coefficients - Partial Regression Coefficients, Standardized Regression Coefficient, Regression Models with Categorical (i.e., Qualitative) Variables - Interpretation of Regression Coefficients of Categorical Variables, Interaction Variables in Regression Models, Validation of Multiple Regression Model, Coefficient of Multiple Determination (R-Squared), Adjusted R-Squared, Statistical Significance of Individual Variables in Multiple Linear Regression: t-Test, Validation of Overall Regression Model: F-Test, Validation of Portions of a Multiple Linear Regression Model - Partial F-Test, Residual Analysis in Multiple Linear Regression.

Practical Component:

(i) Program to demonstrate Multiple Linear Regression

MODULE 4 : INTRODUCTION TO CLASSIFICATION & CLASSIFICATION ALGORITHMS (12)

What is Classification? General Approach to Classification, k-Nearest Neighbor Algorithm, Logistic Regression, Decision Trees, Naive Bayesian Classifier, Ensemble Methods: Bagging, Boosting and

AdaBoost and XBoost, Random Forests, Advanced Classification Methods: Backpropagation in Multilayer Feed-Forward Neural Networks, Support Vector Machines, Rough Set and Fuzzy Set Approaches, Classification Model Evaluation and Selection: Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, Lift Curves and Gain Curves, ROC Curves, Misclassification Cost Adjustment to Reflect Real-World Concerns, Decision Cost/Benefit Analysis.

Practical Component:

(i) Program to demonstrate k-Nearest Neighbor flowers classification

(ii) Program to demonstrate Decision Tree – ID3 Algorithm

(iii) Program to demonstrate Naïve- Bayes Classifier

(iv) Program to demonstrate Back-Propagation Algorithm

MODULE 5: INTRODUCTION TO CLUSTER ANALYSIS & CLUSTERING METHODS (12)

The Clustering Task and the Requirements for Cluster Analysis , Overview of Some Basic Clustering Methods, Hierarchical Methods: Agglomerate versus Divisive Hierarchical Clustering, Distance Measures, Probabilistic Hierarchical Clustering, Multiphase Hierarchical Clustering Using Clustering Feature Trees, Partitioning Methods: k-Means Clustering, k-Medoids Clustering, Density-Based Clustering: DBSCAN - Density-Based Clustering Based on Connected Regions with High Density, Measuring Clustering Goodness.

Practical Component:

(i) Program to demonstrate k-means clustering algorithm

(ii) Program to demonstrate K-Means Clustering Algorithm on Handwritten Dataset

(iii) Program to demonstrate K-Medoid clustering algorithm

(iv) Program to demonstrate DBSCAN clustering algorithm

(v) Program to demonstrate SVM based classification

(vi) Program to demonstrate PCA on face recognition

(vii) Program to demonstrate PCA and LDA on Iris dataset

(viii) Mini Project works shall be given with a batch of four students considering different datasets such as digit dataset, face dataset, flower dataset and micro-array dataset.

TEXT BOOK

1.	Machine Learning (IBM ICE Publications).

SEMESTER-VI

COURSE TITLE	**DEEP LEARNING**			**CREDITS**	**4**
COURSE CODE	**IBA4316**	**COURSE CATEGORY**	**PC**	**L-T-P-C**	**3-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-3**				
CO	**COURSE OUTCOMES**				**PO**

Upon completion of this course, the students will be able to		
1	Understand the fundamentals of Deep Learning	1,2
2	Familiarize with Neural Networks aspects	1,2,5,12
3	Appreciate the concept behind Convolution Neural Networks	1,2,5,12
4	Identify the Generative Deep learning models	1,2,5,12
5	Recognize the tangible applications of ML and Neural Networks.	1,2,5,12

Prerequisites : Machine Learning

MODULE 1: THE FUNDAMENTALS OF DEEP LEARNING (12)

How is Deep Learning different from machine learning? Evolution of AI and ML: Historical Epochs, why is Deep Learning relevant now?

Practical Component:

Hardware requirement:

- **Nvidia GPU**: 8 GB+ GPU RAM
- **CPU RAM**: 32 GB RAM+

Software Environment:

- **OS** - Ubuntu 16x
- **Nvidia CUDA Version** - 10.0
- **Python 2.7.15 and 3.4**+
- **nvidia-smi** (verifiable by typing ***nvidia-smi*** in terminal and checking for GPU memory and running process-information)
  Following Python packages:
- Numpy version 1.15.4
- Tensorflow version 1.12.0
- **Keras** version 2.2.4
- **sklearn** 0.19.2
- **Opencv** (cv2 package) - version 3.4.5

Please note that the hardware requirement specified is the **minimum needed**, and software versions needed are **exact**; installing older and newer versions not recommended.

MODULE 2: THE MATH BEHIND NEURAL NETWORKS AND DEPTHS OF DEEP LEARNING (12)

How does a neural network look like? The matrix magic: Scalars->Vectors->Matrices->Tensors. How they represent real world data? Tensor operations and visualization, Visualizing Deep Learning, Elephant in the room: Gradient Descent: Derivatives and tensors, Stochastic gradient descents, Back propagation: our very own chain rule of differentiation!

Layers: the atoms of DL, Models: the molecules of DL, Loss functions, Optimizers, Activation Functions: Application and types, Deep Learning in Action: This or that: binary classification, Which one is it? Multi-class classification and Regression. Does your network memorize? Overfitting and Under fitting.

Practical Component:

(i) Print Dimensions of dataset
(ii) Calculation of Accuracy Values
(iii) Compose Matrix Shape and Tensor Shape
(iv) Accessing and Manipulation of tensors
(v) Understand the mechanism of practically training a binary classifier

MODULE 3: CONVOLUTIONAL NEURAL NETWORKS AND RECURRENT NEURAL NETWORKS (RNN) (12)

Convolution operation, Max pooling, Power of CNNs: abstraction across layers, what each convolution layer sees? Reusing popular CNNS & fine-tuning.

Why Recurrent Networks, LSTMs & GRU, Examples of simple RNNs, Complex recurrent neural networks: Overfitting in RNNs, Multi-layer RNNs, Multi-directional RNNs. Real-life examples: One-dimensional sequence processing, CNN+RNN.

Practical Component:

(i) Access and manipulation of tensors
(ii) Regression Data Sampling
(iii) Combat Overfitting
(iv) CNN Training
(v) Reuse a model
(vi) Reusing a part of an existing model

MODULE 4: GENERATIVE DEEP LEARNING (12)

Using LSTMs to synthesize text, Neural Style transfer and applications, Image synthesis with variational auto encoders, Generative Adversarial Networks: What does a GAN look like? Generator, Discriminator, Generator vs Discriminator, Training GANs.

MIMO Deep Learning models, Layers graphs: acyclic and directional, can layers share weights? Hyper parameter tuning, Bag of tricks: ensemble of models.

Practical Component:

(i) Stages of the convnet
(ii) Sequence Classification Problem
(iii) Text-to-Speech synthesis
(iv) Text Generation
(v) Automatic Image Captioning with Pytorch using a pre-trained model
(vi) Automatic Image Captioning with Keras
(vii)Facial Recognition

MODULE 5: TANGENTIAL TOPICS OF ML AND NEURAL NETWORKS (12)

Information-Theoretic Machine Learning, Hebbian Learning, Competitive Learning, Boltzmann Learning, RadialBasis Function Networks.

Practical Component:

(i) Digit Recognition using Softmax Regression

(ii) Hand Movement Using OpenCV
(iii) Implementing a recurrent neural network (RNN) in TensorFlow
(iv) RNN that will generate new “Flower” names

TEXT BOOK

1.	Deep Learning (IBM ICE Publications).

SEMESTER-VII

COURSE TITLE	**COMPUTATIONAL LINGUISTIC AND NATURAL LANGUAGE PROCESSING**			**CREDITS**	**4**
COURSE CODE	**IBA4301**	**COURSE CATEGORY**	**PC**	**L-T-P-C**	**4-0-2- 4**
CIA	**60%**			**ESE**	**40%**
LEARNING LEVEL	**BTL-3**				

CO	**COURSE OUTCOMES**	**PO**
Upon completion of this course, the students will be able to		
1	Recognize the concepts of Natural Language Processing.	1,2
2	Familiarize with Text processing, Lexical Analysis, Syntactic and Semantic Analysis aspects.	1,2,5,12
3	Design and develop NLP with Corpus creation with key focus on Tree Bank annotation.	1,2,3,5,12
4	Appreciate the nuances of Statistical techniques with modern speech recognition.	1,2,5,12
5	Recognize the applications of NLP in various sectors like healthcare and education.	1,2,5,6,12

MODULE 1: CLASSICAL APPROACHES TO NATURAL LANGUAGE PROCESSING (9)

Introduction – Context - Classical Toolkit - Text Preprocessing – Tokenization – Sentence Segmentation - Lexical Analysis – Finite State Morphonology – Finite State Morphology – Paradigm based Lexical Analysis - Syntactic Parsing – Cocke-Kasami-Younger Algorithm – Deductive Parsing – LR Parsing – Constraint-based Grammars – Issues in Parsing - Semantic Analysis – Theories and approaches to Semantic Representation – Fine Grained Lexical Semantic Analysis: Case studies - Natural Language Generation – Components of a Generator – Approaches to Text Planning – Linguistic Component.

Practical Component:

Lab Specifications:

Regarding the lab requirement for CL & NLP, as suggested IBM Watson can be implemented along with Python NLTK. Python is open source as required. Pycharm or Jupyter IDEs can be used for the same.

(i) Print Dimensions of dataset
(ii) Calculation of Accuracy Values
(iii) Compose Matrix Shape and Tensor Shape
(iv) Accessing and Manipulation of tensors

MODULE 2: ANNOTATION, TAGGING AND PARSING (9)

Corpus Size, Representation, Sampling – Data Capture – Corpus Markup and Annotation – Multilingual Corpora – Multimodal Corpora -Corpus Annotation Types - Morphosyntactic Annotation – Treebanks: Syntactic, Semantic, and Discourse Annotation - Process of Building Treebanks - Applications of Treebanks - Searching Treebanks. Fundamental Statistical Techniques: Binary Linear Classification - One-versus-All Method for Multi-Category Classification - Maximum Likelihood Estimation - Generative and Discriminative Models - Mixture Model and EM - Sequence Prediction Models. Part-of-Speech Tagging: General Framework – POS Tagging Approaches – Other Statistical and Machine Learning Approaches. Statistical Parsing: Basics - Probabilistic Context-Free Grammars - Generative Models - Discriminative Models - Beyond Supervised Parsing.

Practical Component:

(i) Understand the mechanism of practically training a binary classifier
(ii) Access and manipulation of tensors
(iii) Regression Data Sampling
(iv) Combat Overfitting
(v) CNN Training

MODULE 3: MULTIWORD EXPRESSIONS, WEB DISTANCE AND WORD SIMILARITY, WORD SENSE DISAMBIGUATION (9)

Multiword Expressions: Linguistic Properties of MWEs – Types of MWEs – MWE Classification – Research Issues - Methods of Word Similarity – Normalized Web Distance Method – Kolmogorov Complexity – Information Distance – Normalized Web Distance – Applications – Word Sense Inventories and Problem Characteristics – Applications of Word Sense Disambiguation – Approaches to Sense Disambiguation: Supervised, Lightly Supervised and Unsupervised.

Practical Component:

(i) Reuse a model
(ii) Reusing a part of an existing model
(iii) Stages of the convnet
(iv) Sequence Classification Problem
(v) Text-to-Speech synthesis
(vi) Text Generation

MODULE 4: SPEECH RECOGNITION, ALIGNMENT, STATISTICAL MACHINE TRANSLATION (9)

Modern Speech Recognition: Architectural Components – Historical Developments – Speech Recognition Applications – Technical Challenges and Future Research Directions – Alignment: Basics – Sentence Alignment – Character, Word, Phrase Alignment – Structure and Tree Alignment – Biparsing and ITG Tree Alignment – Statistical Machine Translation: Approaches – Language Models – Parallel Corpora – Word Alignment – Phrase Library – Translation Models – Search Strategies – Research Areas.

Practical Component:

(i) Automatic Image Captioning with Pytorch using a pre-trained model
(ii) Automatic Image Captioning with Keras
(iii) Facial Recognition
(iv) Digit Recognition using Softmax Regression

MODULE 5: APPLICATIONS **(9)**

Information Retrieval – Indexing – IR Models – Evaluation and Failure Analysis – Natural Language Processing and Information Retrieval – Question Answering – Generic Question Answering System – Evaluation of Question Answering system – Multilinguality in Question Answering System – Recent trends and Related Works – Information Extraction – IE with Cascaded Finite State Transducers – Learning based Approaches in IE – Report generation – Emerging Applications of Natural language Generation in Information – Biomedical Text Mining – Sentiment Analysis and Subjectivity.

Practical Component:

(i) Hand Movement Using OpenCV
(ii) Implementing a recurrent neural network (RNN) in TensorFlow
(iii) RNN that will generate new “Flower” names

TEXT BOOK

1.	Introduction to Linguistics and Natural Language Processing (IBM ICE Publications).