SEMESTER - IV

COMPUTER VISION

THEORY

1. Course Description

Programme:	**M.Sc. Data Science**	**Max. Hours: 45**
Course Code:	**P24/CDS/DSE/402**	**Hours per week: 3**
Course Type:	**DISCIPLINE SPECIFIC ELECTIVE**	**Max. Marks: 100**
No. of credits:	**3**	

2. Course Objectives

- To introduce students, the fundamentals of image formation.
- To introduce students to the major ideas, methods, and techniques of computer vision and pattern recognition.
- To develop an appreciation for various issues in the design of computer vision and object recognition systems

3. Course Outcomes

On completion of the course the student will be able to:

CO1: ***Illustrate*** and ***explain*** image formation and processing essentials.
(Cognitive levels – 2)
CO2: ***Apply*** object and feature detection techniques.
(Cognitive level – 3,4)
CO3*: Choose* and ***Design*** the processing of 3D reconstruction.
(Cognitive levels – 5,6)

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

4.Course Content

MODULE I: INTRODUCTION TO COMPUTER VISION, IMAGE FORMATION AND PROCESSING **(15 Hrs.)**

Introduction to computer vision, Image formation: Geometric primitives and transformations, Photometric image formation, The digital camera.Image processing: Point operators, Linear filtering, Non-linear filtering Geometric transformations.

MODULE II:RECOGNITION AND FEATURE DETECTION **(15 Hrs.)**

Recognition: Instance recognition, Image classification, Object detection. Feature detection and matching: Points and patches, Edges and contours, Contour tracking, Lines and vanishing points, Segmentation.

MODULE III: 3D RECONSTRUCTION **(15 Hrs.)**

3D reconstruction: Shape from X, 3D scanning, Surface representations, Point-based representations, Volumetric representations, Model-based reconstruction, Recovering texture maps and albedos.

5. References

1. Computer Vision: Algorithms and Applications 2nd Edition, Richard Szeliski, © 2022 Springer.
2. Computer Vision - A modern approach, by D. Forsyth and J. Ponce, Prentice Hall Robot Vision, by B. K. P. Horn, McGraw-Hill.
3. Haralick& Shapiro, "Computer and Robot Vision", Vol II.
4. G_erardMedioni and Sing Bing Kang "Emerging topics in computer vision".
5. Emanuele Trucco and Alessandro Verri "Introductory Techniques for 3-D Computer Vision", Prentice Hall, 1998.
6. Olivier Faugeras, "Three-Dimensional Computer Vision", The MIT Press, 1993.

PROFESSOR
Department of Computer Scien Engineering
University College of En ring (A)
Osmania Univ
Hyderabad-50

6.Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global Development	Computer Vision serves as a highly accurate pattern recognition tool, enhancing facial recognition, signature analysis, video surveillance, and detection of suspicious activities.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1 and 2	Applying image processing, object detection and feature detection techniques.
EMP	Modules 3	Testing and evaluating 3D reconstruction models.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.
P. V. S

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz
3.	Problem solving	Identifying, formulating, and addressing the detection and processing of images, objects using image processing, object detection and 3D reconstruction techniques.

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 2 – Lab	

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007

b) Model Question Paper- End Semester Exam

COMPUTER VISION

MODEL QUESTION PAPER
THEORY

Course Code: P24/CDS/DSE/402 **Max Marks: 60**
Credits: 3 **Time: 2 ½ Hrs.**

I: Answer the following **3 x 12 = 36**

1. Define computer vision. Explain in detail about Geometric primitives and Transformations of 2D and 3D.

OR

2. Discuss image processing. Explain in detail about point operators.

3. Explain in detail about object detection.

OR

4. Define segmentation. Explain in detail various procedures of segmentation.
5. Explain in detail about 3D scanning with an application.

OR

6. Describe in detail Model based reconstruction.

II. Answer any Four: **4 x 6 = 24**

7. Write a short note on digital camera processing steps.
8. Discuss linear filtering.
9. Describe briefly edges and contours.
10. Explain Snakes and scissors in contour tracking.
11. Write a short note on Surface interpolation.
12. Briefly explain volumetric representations.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

c)Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

NOTE: From
Section-A any 3 questions can be answered (INTERNAL CHOICE).
Section-B any 4 questions can be answered. (EXTERNAL CHOICE)

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(Hrs.)
1	1	2	15
2	2	3,4	15
3	2	5,6	15

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

SEMESTER III

COMPUTER VISION PRACTICAL

1.Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSE/402/P
Course Type: DSE
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 50

2.Course Objective

To explore various techniques of image processing, object detection, feature detection and 3D reconstruction.

3.Course Outcomes

CO1: Exploring image formation and image processing techniques.
CO2: Applying object detection and 3D reconstruction procedures.

4. Practical Sessions

1. Implementing various basic image processing operations in python/matlab/open-CV: Reading image, writing image, conversion of images, and complement of an image
2. Implement contrast adjustment of an image. Implement Histogram processing and equalization.
3. Show Geometric primitives and transformations.
4. Implement the various low pass and high pass filtering mechanisms.
5. Find points and patches by detecting features.
6. Performing/Implementing image segmentation.
7. Implement surface representations of 3D.
8. Object detection and Recognition on available online image datasets
9. Character or digit classification
10. Face classification.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

MODEL QUESTION PAPER

COMPUTER VISION
PRACTICAL

COURSE CODE: P24/CDS/DSE/402/P

Max.Marks: 50
Exam Duration: 2 Hrs.

Answer any one of the following.

1. Write a program to implement contrast adjustment of an image, Histogram processing and equalization.
2. Write a program to find points and patches by detecting features
3. Write a program to implement Object detection and Recognition on available online image datasets
4. Write a program for classification of Character or digit.
5. Write a program to implement Face classification.

Prepared by	Checked & Verified by	Approved by
Ms.Khalida Tabassum Teaching Faculty	Ms. D. Sowjenya Head of the Department	Dr.Uma Joseph Principal

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

SEMESTER - IV

DEEP LEARNING

THEORY

1. Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/402
Course Type: DISCIPLINE SPECIFIC CORE
No. of credits: 3

Max. Hours: 45
Hours per week: 3
Max. Marks: 100

2. Course Objectives

- To understand the basic theory on deep learning.
- To learn the implementation of problems that can be handled by deep learning using Keras.
- To learn the implementation of neural networks.

3. Course Outcomes

On completion of the course the student will be able to:

CO1: ***Illustrate*** and ***explain*** keras-A neural network API and tensor board- Visualization toolkit.
(Cognitive levels – 2)
CO2: ***Apply*** deep learning with convolutional neural networks and Generative Adversarial Networks.
(Cognitive level – 3,4)
CO3*: Choose* and ***Design*** the word embedding and Recurrent neural networks.
(Cognitive levels – 5,6)

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

4. Course Content

MODULE I: FOUNDATIONS OF NEURAL NETWORKS, WORKING WITH KERAS (15 Hrs.)

Perceptron, The first example of Keras code, Multilayer perceptron, Problems in training the perceptron and a solution, Activation functions, sigmoid, ReLU, One-hot encoding, Defining a simple neural net in Keras, Running a simple Keras net and establishing a baseline, Improving the simple net in Keras with hidden layers and with dropout, Testing different optimizers in Keras, Increasing the number of epochs, Controlling the optimizer learning rate, Increasing the number of internal hidden neurons, Increasing the size of batch computation, Summarizing the experiments run for recognizing handwritten charts, Adopting regularization for avoiding overfitting Hyperparameters tuning, Predicting output, A practical overview of backpropagation, Towards a deep learning approach.Installing and Configuring Keras, Keras API, Keras architecture, tensor, Composing models in Keras, Overview of predefined neural network layers, predefined activation functions, metrics, optimizers. Saving and loading the weights and the architecture of a model, Checkpointing, Using TensorBoard and Keras, Quiver and Keras.

MODULE II: DEEP LEARNING WITH CONVOLUTIONAL NETWORKS, GENERATIVE ADVERSARIAL NETWORKS (15 Hrs.)

Deep convolutional neural network, Local receptive fields, Shared weights and bias, Pooling layers, Max and Average pooling, ConvNets, LeNet and LeNet code in Keras, Understanding the power of deep learning, Recognizing and Improving the CIFAR-10 images with deep learning, Very deep convolutional networks for large-scale image recognition, Recognizing cats with a VGG-16 net, Utilizing Keras built-in VGG-16 net module. Overview of Generative Adversarial Networks and GAN applications, Deep convolutional generative adversarial networks, Keras adversarial GANs for forging MNIST and CIFAR, WaveNet.

MODULE III: WORD EMBEDDINGS AND RECURRENT NEURAL NETWORK (15 Hrs.)

Distributed representations, word2vec, skip-gram and CBOW word2vec models, Extracting word2vec embeddings, Using third-party implementations of word2vec, Exploring GloVe, Using pre-trained embeddings and learning from scratch, Fine-tuning learned embeddings from word2vec and GloVe, Look up embeddings. Recurrent Neural Network: SimpleRNN cells, SimpleRNN with Keras, RNN topologies, Vanishing and exploding gradients, Long short term memory, LSTM with Keras, sentiment analysis, Gated recurrent unit, GRU with Keras, POS tagging, Bidirectional and Stateful RNNs.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

5. References

1. Deep Learning with Keras, by Antonio Gulli, Sujit Pal, Packt Publishing, 2017..
2. Deep Learning by Ian Goodfellow, Yoshua Bengio and Aaron Courville published by MIT Press.
3. Neural Networks and Deep Learning: A Textbook, by Charu C. Aggarwal, Springer.
4. Deep Learning (The MIT Press Essential Knowledge series) by John D. Kelleher
5. Deep Learning From Scratch: Building with Python from First Principles by Seth Weidman, O`Reilley.

6.Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	**Relevance**
Global Development	Deep learning has remarkable success across various domains, including computer vision, natural language processing, speech recognition, and reinforcement learning etc.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	**Syllabus Content**	**Description of Activity**
SD	Modules 1 and 2	Applying artificial neural networks using sequential API, CNN, GAN
EMP	Modules 3	Testing and evaluating Word Embeddings and RNN.

PROFESSOR
Department of Computer Science & Engine
University College of Engineering (A)
Osmania University
Hyderabad
P.V. Sridevi

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz
3.	Problem solving	Image Classification, NLP tasks etc.

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 2 – Lab	

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007

b) Model Question Paper- End Semester Exam

DEEP LEARNING

MODEL QUESTION PAPER
THEORY

Course Code: P24/CDS/DSC/402
Credits: 3

Max Marks: 60
Time: 2 ½ Hrs.

I: Answer the following **3 x 12 = 36**

1. Describe in detail about Multilayer perceptrons.

OR

2. Define Keras. Explain in detail about Keras API and architecture.

3. Define the power of deep learning.Discuss Recognizing and Improving the CIFAR-10 images with deep learning,

OR

4. Describe about Deep convolutional generative adversarial networks Numerical perceptron.
5. Explain in detail about Extracting word2vec embeddings and Using third-party Implementations of word2vec.

OR

6. Discuss the Recurrent Neural Network and explain about RNN topologies

II. Answer any Four: **4 x 6 = 24**

7. Discuss about sigmoid, ReLU.
8. Write a short note on TensorBoard.
9. Write a short note on ConvNets, LeNet and LeNet code in Keras.
10. Discuss about WaveNet.
11. Explain about CBOW word2vec models.
12. Write about Stateful RNNs.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

c)Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

NOTE: From
Section-A any 3 questions can be answered (INTERNAL CHOICE).
Section-B any 4 questions can be answered. (EXTERNAL CHOICE)

9. CO-PO Mapping

CO	PO	Cognitive Level	Class room sessions(hrs)
1	1	2	15
2	2	3,4	15
3	2	5,6	15

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

SEMESTER IV

DEEP LEARNING PRACTICAL

1.Course Description

Programme: M.Sc. Data Science
Course Code: P24/CDS/DSC/402/P
Course Type: DSC
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 50

2.Course Objective

To install and customize Keras to explore and implement deep learning mechanisms like generative adversarial networks, recurrent neural networks.

3.Course Outcomes

CO1: Exploring Keras for Generative adversarial networks and convolutional networks.
CO2: Applying Keras API for Word embeddings and RNN.

4. Practical Sessions

1. Neural Networks, Working with Keras:
   a. Running a simple Keras net and establishing a baseline.
   b. Installing and Configuring Keras.

2. Deep Learning with ConvNets, Generative Adversarial Networks:
   a. Recognizing and Improving the images with deep learning
   b. working with Keras adversarial Generative Adversarial Networks

3. Word Embeddings and Recurrent Neural Network:
   a. Extracting word2vec embeddings
   b. Implementing Simple Recurrent Neural Networks with Keras

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

MODEL QUESTION PAPER

DEEP LEARNING
PRACTICAL

COURSE CODE: P24/CDS/DSC/402/P

Max.Marks: 50
Exam Duration: 2 Hrs.

Answer any one of the following.

1. Implement handwritten digit recognition using MNIST dataset with ANN.
2. Write a program for image recognition of CIFAR-10 Image dataset using CNN.
3. Implement Generative Adversarial Networks with keras.
4. write a program representing Word2Vec.
5. Write a program on RNN for finding sequences of words.

Prepared by	Checked & Verified by	Approved by
Ms.Khalida Tabassum Teaching Faculty	Ms. D. Sowjenya Head of the Department	Dr.Uma Joseph Principal

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

SEMESTER -IV

INFORMATION SECURITY

THEORY

1. Course Description

Programme: M.Sc. Data.Science **Max. Hours: 45**
Course Code:P24/CDS/DSE/401 **Hours per week: 3**
Course Type: DISCIPLINE SPECIFIC ELECTIVE **Max. Marks: 100**
No. of credits: 3

2. Course Objectives

- Learn and understand fundamentals of cryptography, network security threats, security services, countermeasures, vulnerability, analysis of network security and its application to network security.
- Acquire & Identify background on hash functions; authentication; firewalls; intrusion detection techniques.
- Apply methods for authentication, access control, intrusion detection and prevention.

3. Course Outcomes

On completion of the course the student will be able to:

CO1: ***Illustrate*** and ***explain*** basic computer concepts of the risks faced by computer systems and networks and analyse security problems. (Cognitive levels – 2)
CO2: ***Apply*** security mechanisms to protect computer systems and networks (Cognitive level – 3,4)
CO3*: Choose* and ***Design*** cryptography algorithms and protocols to achieve computer security (Cognitive levels – 5,6)

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

4. Course Content

MODULE I: INTRODUCTION TO INFORMATION SECURITY (15 Hrs)

Introduction to Information Security: Attacks, Vulnerability, Security Goals, Security Services and mechanisms. Conventional Cryptographic Techniques: Conventional substitution and transposition ciphers, One-time Pad, Block cipher and Stream Cipher.

MODULE II: CRYPTOGRAPHIC TECHNIQUES & AUTHENTICATION (15 Hrs)

Symmetric and Asymmetric Cryptographic Techniques: DES, AES, RSA algorithms.
Authentication and Digital Signatures- Use of Cryptography for authentication, Secure Hash function.Program Security: Non malicious Program errors – Buffer overflow, Incomplete mediation, Time-of-check to Time-of- use Errors, Viruses, Man-in-the-middle attacks.

MODULE III: SECURITY AND FIREWALLS (15 Hrs)

Security in Networks: Threats in networks, Network Security Controls – Architecture, Encryption, Content Integrity, Strong Authentication, Access Controls, Wireless Security, Firewalls – Design and Types of Firewalls, Personal Firewalls.

5. References

1. Security in Computing, Fourth Edition, by Charles P. Pfleeger, Pearson Education, 2015.
2. Cryptography And Network Security Principles and Practice, Fifth Edition, William Stallings, Copyright @2011, 2006 Pearson Education, Inc
3. Modern Cryptography: Theory and Practice, by Wenbo Mao, Prentice Hall.
4. Network Security Essentials: Applications and Standards, by William Stallings. Prentice Hall.

6.Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global Development	Information Security Information security is essential for ensuring that critical business operations can continue in the event of a security breach or cyber-attack. It safeguards all types of data against theft and loss. Sensitive data, protected health information (PHI and is widely employed in everything from system software to game development.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1 and 2	Security protocols to help secure an organisation's data. To protect organizations from vulnerabilities and attacks.
EMP	Modules 3	Applying skills in security, and implementing code. Protecting the organization from network threats.

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz, Group Discussion
3.	Problem solving	Cryptography algorithm implementation

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 2 – Lab	

b) Model Question Paper- End Semester Exam

INFORMATION SECURITY

MODEL QUESTION PAPER

THEORY

Course Code: P24/CDS/DSE/401 **Max Marks: 60**
Credits: 3 **Time: 2 ½ Hrs.**

I: Answer the following **3 x 12 = 36**

1. What Discuss about various Security attacks in detail.

OR

2. List the various substitution techniques. Explain any one technique with an example.

3.What is DES and explain the procedure of encryption in DES.

OR

4. Explain the Time-of-check to Time-of- use Errors of two numbers.

5. Explain Encryption in detail.

OR

6.Define a firewall and describe characteristics, types of firewalls in detail.

II. Answer any Four: **4 x 6 = 24**

5. Explain Vulnerability.
6. Explain the Stream Cipher technique.
7. Write a short note on Buffer overflow.
8. Write a short note on secure hash function?
9. Give an account on threat controls.
10. What are threats? Explain different type of threats.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

c)Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

NOTE: From
Section-A any 3 questions can be answered (INTERNAL CHOICE).
Section-B any 4 questions can be answered. (EXTERNAL CHOICE)

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(Hrs)
1	1	2	15
2	2	3,4	15
3	2	5,6	15

PROFESSOR
Department of Computer Science & Engg.
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

SEMESTER IV

INFORMATION SECURITY
PRACTICAL

1.Course Description

Programme: M.Sc.Data.Science
Course Code: P24/CDS/DSE/401/P
Course Type: DSE
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 50

2.Course Objective

To understand the concepts of information security in a local area network and the Internet.
To apply the security algorithms.

3.Course Outcomes

CO1: Exploring the digital signatures and authentication protocols to find practical solutions to security issues.
CO2: Applying To write programs that use cryptography algorithms and protocols to achieve computer security.

4. Practical Sessions

1. Implement Caesar Cipher.
2. Implement Affine Cipher.
3. Implement Playfair Cipher with a key entered by the user.
4. Implement polyalphabetic Cipher.
5. Implement AutoKey Cipher.
6. Implement Hill Cipher.
7. Implement Rail fence technique.
8. Implement Simple Columnar Transposition technique.
9. Implement Simple RSA Algorithm with small numbers.
10. Implement Simplified DES

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

MODEL QUESTION PAPER

INFORMATION SECURITY
PRACTICAL

COURSE CODE: P24/CDS/DSE/401/P

Max.Marks: 50
Exam Duration: 2 Hrs.

Answer any one of the following.

1. Write a program to implement Caesar Cipher.
2. Write a program to implement Playfair Cipher with a key entered by the user.
4. Write a program to implement polyalphabetic Cipher.
5. Write a program to implement AutoKey Cipher.
6. Write a program to implement simplified DES.

Prepared by	Checked & Verified by	Approved by
Ms. B. Jyothi P.G. Coordinator	Ms. D. Sowjenya Head of the Department	Dr. Uma Joseph Principal

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007

SEMESTER - IV

NATURAL LANGUAGE PROCESSING
THEORY

1. Course Description

Programme: M.Sc.Data. Science
Course Code: P24/CDS/DSC/401
Course Type: DISCIPLINE SPECIFIC CORE
No. of credits: 3

Max. Hours: 45
Hours per week: 3
Max. Marks: 100

2. Course Objectives

- To define and recognize the basic concepts and principles of natural language processing
- To demonstrate ability in text processing techniques
- To design a computational model of natural language

3. Course Outcomes

On completion of the course the student will be able to:

CO1: Apply Python and NLTK to access and process text (Cognitive Level – 3)
CO2: Extract meaningful information from text data using NLP processing (Cognitive Level – 4)
CO3: Evaluate text classification and apply NLTK to perform syntactic analysis and parsing on sentences (Cognitive Level - 5)

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

4. Course Content

MODULE I: NLP and NLTK OVERVIEW (15 Hrs.)

Language Processing and Python - texts and words, texts as lists of words, simple statistics, making decisions and taking control, automatic natural language understanding; accessing text corpora and lexical resources - accessing text corpora, conditional frequency distributions, lexical resources, WordNet

MODULE II: NLP STAGES (15 Hrs.)

Processing raw text: accessing text from the web and from disk, strings, text processing with Unicode, regular expressions for detecting word patterns, useful applications of regular expressions, normalizing text, regular expressions for tokenizing text, segmentation, formatting; categorizing and tagging words - using a tagger, tagged corpora, mapping words to properties using python dictionaries, automatic tagging, n-gram tagging, transformation-based tagging, determining the category of a word

MODULE III: TEXT CLASSIFICATION AND CFG (15 Hrs.)

Learning to Classify Text - supervised classification, evaluation, Naive Bayes classifiers extracting information from text - information extraction, chunking, developing and evaluating chunkers, recursion in linguistic structure, named entity recognition, relation extraction. analyzing sentence structure – basic grammar, use of syntax, context-free grammar, parsing with context-free grammar

5. References

1. Natural Language Processing using Python by Steven Bird, Ewan Klien, EdwardLoper, 1 st edition, O'reilly Publications,2009.
2. Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python, Hobson Lane, Cole Howard, Hannes Hapke, Manning, 2019
3. Speech and Language Processing- Daniel Jurafsky, James H Martin, 2nd edition, PHI, 2008.
4. Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning using Python, Akshay Kulkarni, Adarsha Shivananda,. Apress, 2019

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.
P.V. Sridhar

6.Syllabus Focus

a) Relevance to Local, Regional, National and Global Development Needs

Local /Regional/National /Global Development Needs	Relevance
Global Development	Natural Language Processing (NLP) has significant implications on a global scale, influencing various aspects of communication, information processing, and cross-cultural understanding.

b) Components on Skill Development/Entrepreneurship Development/Employability

SD/ED/EMP	Syllabus Content	Description of Activity
SD	Modules 1 and 2	Design solutions to given problems using NLP algorithms
EMP	Modules 3	Identify document similarity using classification

7. Pedagogy

S. No	Student Centric Methods Adopted	Type / Description of Activity
1.	Participative	Seminars
2.	Experimental	Quiz
3.	Problem solving	Programs

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

P. V. Sridhar

8. Course Assessment Plan

a) Weightage of Marks in Continuous Internal Assessments and End Semester Examination

CO	Continuous Internal Assessments CIA -40%	End Semester Examination-60%
CO1	CIA 1 – Written Exam	Written Exam
CO2	CIA 2 – Assignment	
CO3	CIA 3 – Lab Exercise	

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500007.

b) Model Question Paper- End Semester Exam

NATURAL LANGUAGE PROCESSING

MODEL QUESTION PAPER
THEORY

Course Code: P24/CDS/DSC/401 **Max Marks: 60**
Credits: 3 **Time: 2 ½ Hrs.**

I: Answer the following: **3 x 12 = 36**

1. Explain about the different language understanding techniques using appropriate examples

OR

2. How is conditional frequency distribution useful for text processing? Explain with an example

3. Define the various stages of NLP and elaborate their significance.

OR

4. Explain the use of regular expressions for tokenizing text. Use appropriate examples

5. What is meant by parts of speech tagging? Explain the process of entity name normalization.

OR

6. Explain parsing with context-free grammar with an appropriate example.

II. Answer any Four: **4 x 6 = 24**

1. Write a short note on text corpora.
2. What is WordNeT?
3. Explain the process of normalizing text with an example
4. What are regular expressions?
5. Write a short note on latent semantic analysis.
6. Explain the process of chunking
7. Explain supervised classification.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

c)Question Paper Blueprint

Modules	Hours Allotted in the Syllabus	COs Addressed	Section A (No. of Questions)	Total Marks	Section B (No. of Questions)	Total Marks
I	15	1	2	12	2	12
II	15	2	2	12	2	6
III	15	3	2	12	2	6

9. CO-PO Mapping

CO	PO	Cognitive Level	Classroom sessions(Hrs)
1	1	3	15
2	2	4	15
3	2	5	15

PROFESSOR
Department of Computer Science
University College of Engineering
Osmania University
Hyderabad-500 007

NATURAL LANGUAGE PROCESSING PRACTICAL

2. Course Description

Programme: M.Sc. Data.Science
Course Code: P24/CDS/DSC/401/P
Course Type: DISCIPLINE SPECIFIC CORE
No. of credits: 2

Max. Hours: 60
Hours per week: 4
Max. Marks: 50

2. Course Objectives

- Develop NLP algorithms using NLTK, scikit

3. Course Outcomes

On completion of the course the student will be able to:

CO1: Identify and apply NLP concepts to solve real world problems.
CO2: Use NLTK, scikit, and other relevant Python libraries to generate accurate solutions for NLP domain problems

4. Practical Sessions

1. Working with Regular expressions
2. Performing text preprocessing – tokenization, validation, stemming
3. Performing morphological analysis
4. Implementing n-gram extraction
5. Implementing POS tagging
6. Implement chunking to extract noun phrases
7. Identify semantic relationships between the words from the given text using WordNet dictionary
8. Perform Named Entity Recognition (NER) on given text
9. Mini project (Retrieval based chatbot/Search based chatbot)

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

MODEL QUESTION PAPER

NATURAL LANGUAGE PROCESSING
PRACTICAL

COURSE CODE: P24/CDS/DSC/401/P

Max.Marks: 50
Exam Duration: 2 hrs

Answer any one of the following.

1. Write a program to perform POS tagging.
2. Write a program to perform morphological analysis.
3. Write a program to implement bi-gram extraction.
4. Write a program to implement chunks on given text to extract noun phrases.
5. Write a program to perform NER on given text.

Prepared by	Checked & Verified by	Approved by
Ms. Padmashree Josyula (Faculty, Department of Computer Science)	**Ms. D.Sowjenya (HoD, Dept of Computer Science)**	**Dr. Uma Joseph (Principal, St.Francis College for Women)**

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007

SEMESTER-IV
PROJECT

1. Course Description

Programme: M.Sc. Data.Science
Course Code: P24/DSC/PRJ/401
Course Type: PRJ
No. of Credits: 5

Max. Hours:90
Hours per week: 5
Max Marks: 150

2. Course Objectives

1. To enable students to gain hands on experience
2. To enable students to create awareness of the characteristics of domain areas where data analytics can be applied for creating a solution for a problem.

3. Course Outcomes

1: Demonstrate an ability to work in teams and manage the conduct of the research study.

2: Students gain analytical skills and the ability to probe into new areas of study of research.

4. Evaluation

INTERNAL EVALUATION(Marks): **50**

EXTERNAL EVALUATION(Marks): **100**

Prepared by Course Teacher]	Checked & Verified by HoD / Programme Coordinator	Approved bythe Principal
B. Jyothi PG Coordinator, Dept. of Computer Science	D.Sowjenya, Head, Department of Computer Science	Dr. Uma Joseph Principal, St. Francis College for Women

PROFESSOR
Department of Computer Science & Engineering (A)
University College of Engineering
Osmania University
Hyderabad-500 007.