



Sri Indu

College of Engineering & Technology

UGC Autonomous Institution

Recognized under 2(f) & 12(B) of UGC Act 1956,

NAAC, Approved by AICTE &

Permanently Affiliated to JNTUH



NAAC
NATIONAL ASSESSMENT AND
ACCREDITATION COUNCIL



HANDOUT

III Year II Semester

DEPARTMENT OF COMPUTER SCIENCE &
ENGINEERING (AIML)

ACADEMIC YEAR 2025-2026



SRI INDU COLLEGE OF ENGINEERING & TECHNOLOGY

(An Autonomous Institution under UGC, New Delhi)

(Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi and Accredited by NBA, NAAC)
 Sheriguda Village, Ibrahimpatnam Mandal, RangaReddy Dist.-501510

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING (AIML)

HANDOUT-INDEX

S. No	Contents
1	Vision, Mission, PEOs, POs, PSOs & Cos
2	Institution Academic Calendar
3	Department Academic Calendar
4	Subject wise
i)	Syllabus Copy
ii)	Lesson Plan
iii)	Question Bank
iv)	End Examination Questions (Model Paper)
v)	Mid-1 & Mid-2 Questions (Model Paper)



SRIINDUCOLLEGE OF ENGINEERING & TECHNOLOGY

(An Autonomous Institution under UGC, New Delhi)

(Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi and Accredited by NBA, NAAC)

Sheriguda Village, Ibrahimpatnam Mandal, Ranga Reddy Dist. - 501510

VISION OF THE INSTITUTE

To be a premier institution in engineering & technology and management for competency, values and social consciousness.

MISSION OF THE INSTITUTE

IM₁: Provide high quality academic programs, training activities and research facilities.

IM₂: Promote continuous industry-institute interaction aimed at promoting employability, entrepreneurship, leadership and research aptitude among stake holders.

IM₃: Contribute to the economic and technological development of the region, state and nation.

VISION OF THE DEPARTMENT

To develop competent engineers in the domain of Artificial Intelligence and Machine

Learning for noteworthy contributions to the society.

MISSION OF THE DEPARTMENT

The Department has following Missions:

DM1: To strengthen academic collaborations for better exposure.

DM2: Promote professional environment to imbibe ethical values and entrepreneurial skills

DM3: Encourage research and development by emphasizing on innovation

DM4: Facilitate Industry-Institute collaboration for mutual benefit

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

PEO1: Generate contributions towards advancements in Artificial Intelligence and Machine Learning

PEO2: Promote design, research and implementation of products through strong communication skills, leadership and entrepreneurial skills.

PEO3: Apply basic principles and practices of AIML to successfully complete software related projects to meet customer business objectives and/or productively engage in research. [Typetext]

Program Outcomes (POs)

PO1	Engineering Knowledge: Apply knowledge of mathematics, natural science, computing, engineering fundamentals and an engineering specialization as specified in WK1 to WK4 respectively to develop to the solution of complex engineering problems.
PO2	Problem Analysis: Identify, formulate, review research literature and analyze complex engineering problems reaching substantiated conclusions with consideration for sustainable development. (WK1 to WK4)
PO3	Design/Development of Solutions: Design creative solutions for complex engineering problems and design/develop systems/components/processes to meet identified needs with consideration for the public health and safety, whole-life cost, net zero carbon, culture, society and environment as required. (WK5)
PO4	Conduct Investigations of Complex Problems: Conduct investigations of complex engineering problems using research-based knowledge including design of experiments, modeling, analysis & interpretation of data to provide valid conclusions. (WK8).
PO5	Engineering Tool Usage: Create, select and apply appropriate techniques, resources and modern engineering & IT tools, including prediction and modeling recognizing their limitations to solve complex engineering problems. (WK2 and WK6)
PO6	The Engineer and The World: Analyze and evaluate societal and environmental aspects while solving complex engineering problems for its impact on sustainability with reference to economy, health, safety, legal framework, culture and environment. (WK1, WK5, and WK7).
PO7	Ethics: Apply ethical principles and commit to professional ethics, human values, diversity and inclusion; adhere to national & international laws. (WK9)
PO8	Individual and Collaborative Team work: Function effectively as an individual, and as a member or leader in diverse/multi-disciplinary teams.
PO9	Communication: Communicate effectively and inclusively within the engineering community and society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations considering cultural, language, and learning differences
PO10	Project Management and Finance: Apply knowledge and understanding of engineering management principles and economic decision-making and apply these to one's own work, as a member and leader in a team, and to manage projects and in multidisciplinary environments.
PO11	Life-Long Learning: Recognize the need for, and have the preparation and ability for i) independent and life-long learning ii) adaptability to new and emerging technologies and iii) critical thinking in the broadest context of technological change. (WK8)

	Program specific Outcomes (PSOs)
PSO1	Register mathematical methodology to crack problems using suitable data structures. Register mathematical methodology to crack problems using suitable data structures.
PSO2	Competence to design and develop software for web based and mobiles android sunder real world environment
PSO3	Skill to design the algorithms for machine learning, data compression can be used in different applications.

COURSE OUTCOMES (CO's)

Academic Year: 2025-26

Class: IIIYEAR-II SEM.

Course Name: NATURAL LANGUAGE PROCESSING (R22CSM3213)

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

Course Outcomes(COs)	
C3241.1	Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
C3241.2	Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
C3241.3	Manipulate probabilities, construct statistical models over strings and trees.
C3241.4	Estimate parameters using supervised and unsupervised training methods.
C3241.5	Design, implement, and analyze NLP algorithms.
C3241.6	Design different language modeling Techniques.

COURSE ARTICULATION MATRIX

CO	PO												PSO		
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
C3241.1	3	3	1	-	1	-	-	-	-	-	-	-	-	-	-
C3241.2	3	3	1	-	3	-	-	-	-	-	-	1	1	1	2
C3241.3	3	1	3	-	3	-	-	-	-	-	-	1	1	1	2
C3241.4	3	3	3	-	3	-	-	-	2	-	-	1	-	1	2
C3241.5	3	3	1	-	3	-	-	1	2	-	-	1	-	1	-
C3241.6	3	2	1	-	1	-	-	1	1	-	-	1	-	-	-
C3241	3	2.5	1.6	-	2.5	-	-	0.3	1	-	-	1	0.3	0.6	1

1. Low 2. Medium 3: High

ACADEMIC CALENDAR



SRI INDU COLLEGE OF ENGINEERING & TECHNOLOGY
(An Autonomous Institution under UGC, New Delhi)
Recognized under 2(f) and 12(B) of UGC Act 1956
NBA & NAAC Accredited. Approved by AICTE and Permanently affiliated to JNTUH
Sheriguda (V), Ibrahimpatnam, R.R. Dist, Hyderabad - 501 510

D4

BR-22

Lr.No.SICET/AUTO/DAE/III B.Tech Academic Calendar/29/2025

ANNEXURE -II

B.TECH III-YEAR I-SEM & II-SEM ACADEMIC CALENDAR **(FOR ACADEMIC YEAR : 2025-26)**

Academic Calendar for B.Tech – IIIrd Year Students (2023 - 24 Batch), BR-22 Regulation.

I – Semester

S.No.	EVENT	PERIOD	DURATION
1.	Commencement of class work.	14.07.2025 (Monday)	
2.	1 st Spell of Instructions for covering First Two and a half Units.	14.07.2025 – 13.09.2025	9 Weeks
3.	CRT Classes	15.09.2025 – 07.10.2025	2 Weeks 2 Days
4.	I Mid Term Examinations.	08.10.2025 – 10.10.2025	3 Days
5.	Submission of I Mid Term Examination Marks.	16.10.2025	
6.	2 nd Spell of Instructions for Remaining Two and a half Units.	13.10.2025 – 13.12.2025	9 Weeks
7.	II Mid Term Examinations.	15.12.2025 – 17.12.2025	3 Days
8.	Preparation Holidays and Practical Examinations	18.12.2025 – 27.12.2025	1 Week
9.	Submission of II Mid Term Examination Marks.	23.12.2025	
10.	I Semester End Examinations.	29.12.2025 – 10.01.2026	2 Weeks
Commencement of Class-Work for III B.Tech - II Semester 19.01.2026			

Note: No of Working/instructional days : 90

II – Semester

S.No.	EVENT	PERIOD	DURATION
1.	Commencement of class work.	19.01.2026 (Monday)	
2.	1 st Spell of Instructions for covering First Two and a half Units.	19.01.2026 – 14.03.2026	8 Weeks
3.	I Mid Term Examinations.	16.03.2026 – 18.03.2026	3 Days
4.	Submission of I Mid Term Examination Marks.	24.03.2026	
5.	2 nd Spell of Instructions for Remaining Two and a half Units.	20.03.2026 – 09.05.2026	7 Weeks
6.	Continuation of remaining 2 nd Spell of Instructions	25.05.2026 – 10.06.2026	2 Weeks
7.	II Mid Term Examinations.	11.06.2026 – 13.06.2026	3 Days
8.	Preparation Holidays and Project Evaluation.	15.06.2026 – 20.06.2026	1 Week
9.	Submission of II Mid Term Examination Marks.	18.06.2026	
10.	II Semester End Examinations	22.06.2026 – 04.07.2026	2 Weeks

Note: No of Working/instructional days : 93

ACE

CE

DIRECTOR

PRINCIPAL

Controller of Examination
Sri Indu College of Engineering & Technology
(An Autonomous Institution Under UGC)
Sheriguda (V), Ibrahimpatnam, R.R. Dist-501510

DIRECTOR
Sri Indu College of Engineering & Technology
(An Autonomous Institution Under UGC)
Sheriguda (V), Ibrahimpatnam, R.R. Dist-501510.

PRINCIPAL
Sri Indu College of Engineering & Technology
(An Autonomous Institution Under UGC)
Sheriguda (V), Ibrahimpatnam, R.R. Dist-501510

Sign:

Dr. M.V.S.S. Giridhar
Prof. of CEA; JNTUH Nominee

Sign:

Dr. T. Venu Gopal
Prof. of CSE; JNTUH Nominee

Sign:

Dr. D. Ramesh
Prof. of CSE; JNTUH Nominee

(Copy to DAP&E and Copy to all the Dept. Heads)

DEPARTMENT CALENDAR

SRI INDU COLLEGE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING(AI&ML) DEPARTMENT CALENDAR - 2025-2026(SEMESTER-II)

				Feb-25		Mar-25					
SUNDAY		Dec-25		1		1					Jun-25
MONDAY	1			2		2					1
TUESDAY	2			3	II YR CRT TRAINING	3	II YR CRT TRAINING		Apr-25		2
WEDNESDAY	3			4		4	HOU HOLIDAY	1			3
THURSDAY	4		1	5		5		2	II YR TRAINING	May-25	4
FRIDAY	5	AI CLUB MEETING	2	6	II YR CRT TRAINING	6	II YR CRT TRAINING	3	GOOD FRIDAY HOLIDAY	1	5
SATURDAY	6	DEPARTMENT MEETING	3	7	AI CLUB MEETING	7		4	AI CLUB MEETING	2	6
SUNDAY	7	HOLIDAY	4	8	HOUDAY	8	HOLIDAY	5	HOLIDAY	3	7
MONDAY	8		5	9		9		6		4	8
TUESDAY	9		6	10	II YR CRT TRAINING	10	II YR CRT TRAINING	7	II YR TRAINING	5	9
WEDNESDAY	10		7	11		11		8		6	10
THURSDAY	11		8	12		12		9		7	11
FRIDAY	12		9	13	II YR CRT TRAINING	13	II YR CRT TRAINING	10	II YR TRAINING	8	12
SATURDAY	13		10	14		14		11	DEPARTMENT MEETING	9	13
SUNDAY	14	HOLIDAY	11	15	HOUDAY	15	HOLIDAY	12	HOLIDAY	10	14
MONDAY	15	IV-II COMMENCEMENT WORK	12	16	IV-II MID I EXAM	16	II & III-II MID I EXAM	13		11	15
TUESDAY	16		13	17	IV-II MID I EXAM, II YR TRAINING	17	II & III-II MID I EXAM, II YR TRAINING	14	DRAMBEDHAR JAYANTI HOLIDAY	12	16
WEDNESDAY	17		14	18	IV-II MID I EXAM	18	II & III-II MID I EXAM	15		13	17
THURSDAY	18		15	19	IV-II 2ND SPELL	19		16		14	18
FRIDAY	19		16	20	II YR CRT TRAINING	20	II & III-II MID I SPELL INSTRUCTIONS	17	II YR TRAINING	15	19
SATURDAY	20		17	21		21		18		16	20
SUNDAY	21	HOLIDAY	18	22	HOUDAY	22	HOLIDAY	19	HOLIDAY	17	21
MONDAY	22		19	23		23		20		18	22
TUESDAY	23		20	24	SUBMISSION OF IV-II MID I MARKS, II YR TRAINING	24	SUBMISSION OF III-II MID I MARKS	21	II YR TRAINING	19	23

NATURAL LANGUAGE PROCESSING

SYLLABUS

BR22 – B.TECH. – CSE(ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)

SRI INDU COLLEGE OF ENGINEERING & TECHNOLOGY
(An Autonomous Institution under UGC, New Delhi)

B.Tech. - III Year - II Semester

L	T	P	C
3	0	0	3

(R22CSM3213) NATURAL LANGUAGE PROCESSING

Course Objectives:

- Introduce to some of the problems and solutions of NLP and their relation to linguistics and statistics.

Course Outcomes:

- Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
- Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems
- Able to manipulate probabilities, construct statistical models over strings and trees, and estimate parameters using supervised and unsupervised training methods.
- Able to design, implement, and analyze NLP algorithms
- Able to design different language modeling Techniques.

UNIT - I

Finding the Structure of Words: Words and Their Components, Issues and Challenges, Morphological Models

Finding the Structure of Documents: Introduction, Methods, Complexity of the Approaches, Performances of the Approaches

UNIT - II

Syntax Analysis: Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Parsing Algorithms, Models for Ambiguity Resolution in Parsing, Multilingual Issues

UNIT - III

Semantic Parsing: Introduction, Semantic Interpretation, System Paradigms, Word Sense Systems, Software.

UNIT - IV

Predicate-Argument Structure, Meaning Representation Systems, Software.

UNIT - V

Discourse Processing: Cohension, Reference Resolution, Discourse Cohension and Structure Language Modeling: Introduction, N-Gram Models, Language Model Evaluation, Parameter Estimation, Language Model Adaptation, Types of Language Models, Language-Specific Modeling Problems, Multilingual and Cross lingual Language Modeling

TEXT BOOKS:

1. Multilingual natural Language Processing Applications: From Theory to Practice – Daniel M. Bikel and Imed Zitouni, Pearson Publication.
2. Natural Language Processing and Information Retrieval: Tanvir Siddiqui, U.S. Tiwary.

REFERENCE BOOK:

1. Speech and Natural Language Processing - Daniel Jurafsky & James H Martin, Pearson Publications.

	SRI INDU COLLEGE OF ENGG & TECHLESSONPLAN (Regulation :R22) Department of AIML			Prepared on Rev1: Page:1of3
	Sub.Code &Title	R22CSM3213 & Natural Language Processing		
	AcademicYear:2025-26	Year/Sem	III/II	
	Faculty Name & Designation	M. SWATHI REDDY,Asst.prof., B.PAVAN KUMAR,Asst.Prof		

Unit/Item No.	Topic(s)	Book Reference	Page(s)		Teaching Methodology	Proposed No.of Periods	Actual Date of Handled	CO/RBT
			From	To				
UNIT-I								
I	FINDING THE STRUCTURE OF WORDS					12		
1.1	Finding the Structure of Words Words and Their Components	T1	3	7	Black board	01		CO1/L1
1.2	Issues and Challenges	T1	8	15	Black board	01		CO1/L3
1.3	Morphological Models	T1	15	21	Black board	03		CO1/L2
1.4	Finding the Structure of Documents: Introduction	T1	29	33	Black board	01		CO1/L1
1.5	Methods	T1	33	40	Black board	03		CO1/L2
1.6	Complexity of the Approaches	T1	40	41	Black board	01		CO1/L3
1.7	Performances of the Approaches	T1	41	47	Black board	02		CO1/L3
Review		Signature of the HOD/Coordinator						
Unit/ItemNo.	Topic(s)	Book Reference	Page(s)		Teaching Methodology	Proposed No.of Periods	Actual Date of Handled	CO/RBT
			From	To				
UNIT-II								
II	SYNTAX ANALYSIS					12		
2.1	Syntax Analysis: Introduction	T1	57	57	Black board	01		CO2/L1
2.2	Parsing Natural Language	T1	57	59	Black board	01		CO2/L2
2.3	Treebanks: A Data-Driven Approach to Syntax	T1	59	62	Black board	01		CO2/L1
2.4	Representation of Syntactic Structure	T1	63	70	Black board	02		CO2/L2
2.5	Parsing Algorithms	T1	70	80	Black board	02		CO2/L3
2.6	Models for Ambiguity Resolution in Parsing	T1	80	87	Presentation	03		CO2/L6
2.7	Multilingual Issues	T1	87	92	Black board	02		CO2/L4

		SRI INDU COLLEGE OF ENGG & TECH					LESSONPLAN		Prepared on		
		(Regulation :R22)					Department of AIML		Rev1:		
									Page:2of3		
Sub.Code &Title		R22CSM3213& Natural Language Processing									
AcademicYear:2025-26				Year/Sem			III/II				
Faculty Name &Designation				M. SWATHI REDDY,Asst.prof., B.PAVAN KUMAR,Asst.Prof							
Review			Signature of the HOD/Coordinator								
Unit/ Item no	Topic(s)	Book Reference	Page(s) From To		Teaching Methodology	Proposed No.of Periods	Actual Date of Handled	CO/RBT			
UNIT III											
III	SEMANTIC PARSING					14					
3.1	Semantic parsing: Introduction	T1	97	98	Black board	02		CO3/L2			
3.2	Semantic Interpretation	T1	98	99	Black board	02		CO3/L1			
3.3	Structural Ambiguity	T1	99	101	Black board	01		CO3/L4			
3.4	System Paradigms	T1	101	102	Black board	02		CO3/L3			
3.5	Word Sense	T1	102	104	Black board	02		CO4/L2			
3.6	Resources	T1	104	105	Black board	02		CO4/L4			
3.7	Systems	T1	105	116	Black board	01		CO4/L2			
3.8	Software	T1	116	117	Black board	02		CO4/L5			
Review			Signature of the HOD/Coordinator								
Unit/ Item no	Topic(s)	Book Reference	Page(s) From To		Teaching Methodology	Proposed No.of periods	Actual Date of Handled	CO/RBT			
UNIT IV											
IV	PREDICATE-ARGUMENT STRUCTURE					13					
4.1	Introduction	T1	118	118	Black board	01		CO5/L1			
4.2	Predicate-Argument Structure	T1	118	118	Black board	02		CO5/L2			
4.3	Resources	T1	118	122	Black board	02		CO5/L3			
4.4	Systems, Software	T1	122	147	Black board	03		CO5/L3			

	SRI INDU COLLEGE OF ENGG & TECH LESSON PLAN (Regulation:R22) Department of AIML		
	Sub.Code &Title	R22CSM3213&Natural Language Processing	
	AcademicYear:2025-26	Year/Sem	III/II
	Faculty Name &Designation	M. SWATHI REDDY,Asst.prof., B.PAVAN KUMAR,Asst.Prof	

LIST OF TEXT BOOKS AND REFERENCES

TEXTBOOKS:

- T1. Multilingual natural Language Processing Applications: From Theory to practice-Daniel M.Bikel and imed Zitouni, Pearson Publication.
- T2. Natural Language Processing and Information retrieval: Tanvier Siddiqui, US Tiwary

REFERENCE BOOKS:

1. Speech and Natural Language Processing - Daniel Jurafsky & James H Martin, Pearson Publications

WEB LINK:

<https://www.youtube.com/watch?v=l7IYZLX-Uro>

	SRI INDU COLLEGE OF ENGG & TECHLESSONPLAN (Regulation:R22) Department of AIML		
	Sub.Code &Title	R22CSM3213&NATURAL LANGUAGE PROCESSING	
	AcademicYear:2025-26	Year/Sem	III/II
	Faculty Name & Designation	M. SWATHI REDDY,Asst.prof., B.PAVAN KUMAR,Asst.Prof	

LIST OF POWERPOINT PRESENTATION

S.No	Topic Name	No. Of Slides
1	Finding the Structure of Words and Documents	40 Slides(PPT1)
2	Semantic Analysis	18 Slides(PPT2)

CONTENT BEYOND THE SYLLABUS

S.N O	Topics	Proposed Actions	Date	Resource Person/Mode	POs	PSOs
1	Truth Maintenance Systems	Class room (1Periods)			PO1,PO2	PSO1,PSO2
2	Blackboard Systems	Classroom (1Periods)			PO1,PO2	PSO1,PSO2
3	Fuzzy logic	Classroom (5Periods)			PO1,PO4	PSO4

SELFSTUDYTOPICS			
S.No.	Topics	Books & Journals	Course Outcomes
1	Machine-Learning Paradigms	Saroj Kaushik	CO4,CO5/L3
2	Deep Learning	Yoshua Bengio, and Aaron Courville	CO4,CO5/L3
3	Introduction to Artificial intelligence	NPTEL	CO2/L3

Prepared by	Recommended and Approved by
M. SWATHI REDDY (Signature &Name)	HOD

	SRI INDU COLLEGE OF ENGG & TECH LESSON PLAN (Regulation :R22) Department of AIML		
	Sub.Code & Title	R22CSM3213&NATURAL LANGUAGE PROCESSING	
	Academic Year:2025-26	Year/Sem	III/II
	Faculty Name & Designation	M. SWATHI REDDY,Asst.prof., B.PAVAN KUMAR,Asst.Prof	

QUESTION BANK WITH BLOOMS TAXONOMY LEVEL (BTL)

(1. Remembering 2. Understanding 3. Applying 4. Analyzing 5. Evaluating 6. Creating)

UNIT I			
Multiple Choice Questions		BT Level	Course Outcome
#1C-1	Which of the following is a component of a word's structure? A) Morpheme B) Sentence C) Paragraph D) Phoneme	I	C01
#1C-2	What is the primary function of a morpheme in word formation? A) To change the meaning of a word B) To modify sentence structure C) To provide grammatical tense D) To serve as a part of a larger sentence	IV	C01
#1C-3	Which of the following morphological models focuses on analyzing word forms based on the root and affixes? A) Context-free grammar B) Finite-state morphology C) Dependency grammar D) Syntactic parsing	III	C01
#1C-4	What is one of the main challenges in analyzing the structure of words in languages with rich morphology? A) Lack of vocabulary B) Lack of clear syntax rules C) Ambiguity in word forms due to inflection and derivation D) Too few word variations	I	C01
#1C-5	Which of the following morphological processes involves adding prefixes or suffixes to modify the meaning of a word? A) Compounding B) Derivation C) Inflection D) Clipping	III	C01
#1C-6	In morphology, what does the term "inflection" refer to? A) Changing the tense or case of a word without altering its meaning B) Creating a new word from an existing root C) Combining two words to form a new word D) Changing the syntactic structure of a sentence	II	C01
#1C-7	Which of the following is a major challenge when working with	I	C01

	word morphology in languages with complex inflection systems? A) Identifying the root word B) Parsing syntax C) Handling word ambiguity D) Analyzing sentence meaning		
#1C-8	What does the term “stem” refer to in morphology? A) The plural form of a word B) The central part of a word that carries its primary meaning C) The process of changing a word's tense D) A suffix added to a word	VI	C01
#1C-9	Which type of morphology involves using dictionaries to look up word forms and their meanings? A) Analytic morphology B) Synthetic morphology C) Derivational morphology D) Lexical morphology	II	C01
#1C-10	What is one challenge that arises in agglutinative languages (such as Turkish or Finnish) when processing word structure? A) Identifying individual morphemes due to long compound words B) Lack of consistent word order C) Minimal morphological variation D) Syntactic ambiguity	I	C01
#1C-11	Which of the following is a key step in analyzing the structure of documents in natural language processing? A) Identifying word stems B) Text segmentation C) Sentence parsing D) Morphological analysis	I	C01
#1C-12	In the context of document analysis, what is a major challenge when dealing with large volumes of text data? A) The lack of proper grammar rules B) Ambiguity in sentence structure C) Performance and scalability of processing techniques D) Lack of labeled training data	II	C01
#1C-13	Which method is commonly used to analyze and extract structure from a document? A) Support Vector Machines B) Latent Semantic Analysis (LSA) C) Text segmentation algorithms D) Dependency parsing	III	C01
#1C-14	Which of the following is a significant concern when using machine learning models for document structure analysis? A) Complexity and time required for training B) Lack of reliable documents C) Excessive word repetition D) Uniformity of syntax	I	C01
#1C-15	Which of the following is a key issue in evaluating the performance of document analysis models? A) Determining the stem of words B) Balancing precision and recall in classification tasks C) Identifying synonyms D) Analyzing punctuation use	IV	C01
	Filling The Blank Questions		

1F-1	Morphology is the study of the _____ and _____ of words.	II	C01
1F-2	The smallest unit of meaning in a word is called a _____.	IV	C01
1F-3	_____ morphology refers to the way words are formed from smaller meaningful units like prefixes, suffixes, and roots.	I	C01
1F-4	A _____ morpheme carries a grammatical function, such as tense or number, without changing the word's core meaning.	I	C01
1F -5	One of the key challenges in morphological analysis is handling _____, where one word form can have multiple meanings depending on context.	III	C01
1F -6	The process of breaking down a word into its constituent morphemes is called _____.	I	C01
1F -7	_____ models use a predefined set of rules to generate word forms based on known morphological structures.	I	C01
1F -8	The study of _____ morphemes examines how new words are created by adding prefixes or suffixes to existing words.	II	C01
1F -9	Document structure refers to the _____ in which the content is organized, influencing its clarity and understanding.	VI	C01
1F -10	One common method of analyzing document structure is _____, where a machine identifies key topics based on the frequency and relationships of terms.	I	C01
1F -11	In document analysis, _____ analysis focuses on extracting and analyzing entities or relationships mentioned in the text.	III	C01
1F -12	_____ models are statistical techniques that represent documents as distributions of topics, allowing for topic extraction and categorization.	V	C01
1F -13	The complexity of document analysis approaches often arises from _____, where documents contain large amounts of unstructured data with varied vocabulary.	I	C01
1F -14	One of the main challenges of document classification is the _____ problem, where the same concept can be described using different words or phrases.	II	C01
1F -15	The performance of document analysis approaches is typically evaluated by metrics such as _____ and _____.	II	C01

Match The Following Questions

1M-1		1. Morpheme	A. The smallest unit of meaning in a word		a)		
		2. Root	B. Morphemes that modify the meaning of a word and create new words				
		3. Derivational Morpheme	C. A word's basic form, to which affixes can be added	II		C01	
		4. Inflectional Morpheme	D. Morphemes that alter a word's grammatical function (e.g., tense, number)				
1M-2		1. Document Classification	A. A method of discovering hidden topics within a set of documents based on patterns		b)		
		2. Tokenization	B. Breaking down text into smaller units, such as words or phrases		III		C01
		3. Named Entity Recognition	C. Identifying specific entities like names, dates, or locations within text				
		4. Topic Modeling	D. The process of categorizing documents into predefined classes based on content				

1M-3		1. Rule-Based Model	A. A model that uses a set of states and transitions to process morphological patterns	c)		
		2. Finite State Automaton	B. A model that uses predefined rules to analyze or generate word forms	II	CO1	
		3. Stemming	C. A method of reducing words to their base form by removing suffixes			
		4. Lemmatization	D. A process of reducing a word to its canonical dictionary form			
1M-4		1. Bag-of-Words Model	A. A model that represents documents as a collection of word frequencies without considering word order	d)		
		2. N-Gram Model	B. A technique that finds patterns and relationships between words and their meanings across large datasets			
		3. Latent Semantic Analysis (LSA)	C. A machine learning algorithm that classifies text based on a hyperplane that separates different categories	I	CO1	
		4. Support Vector Machines (SVM)	D. A probabilistic model that predicts the likelihood of word sequences based on the previous N words			
1M-5		1. Ambiguity in Word Structure	A. When a word has multiple meanings depending on the context	e)		
		2. Polysemy	B. The challenge of processing large documents or datasets efficiently	)		
		3. Document Complexity	C. A challenge in analyzing documents due to their varied structures, formats, and content	i) II	CO1	
		4. Scalability	D. The presence of multiple interpretations of a word's form or meaning based on context			

5 MARKS QUESTIONS

1D-1	Differentiate between inflectional and derivational based Morphology.	IV	CO1
1D-2	Explain Token, Lexeme, and Morpheme.	II	CO1
1D-3	Define Morphology. Explain the issues and challenges in morphological processing. (Jan 2024)	II	CO1
1D-4	Differentiate between sentence boundary detection and topic boundary detection with suitable example. (June 2023)	II	CO1
1D-5	Write a summery on Natural Language Processing.	VI	CO1
1D-6	Differentiate between Generative and Discriminative Methods.	IV	CO1
1D-7	Explain various issues and challenges faced in morphology.	I	CO1
1D-8	How the performance of the approaches is evaluated.	V	CO1
1D-9	Differentiate between Dictionary lookup and functional morphology.	IV	CO1
1D-10	Write about the necessary steps involved in natural language processing?	III	CO1
1D-11	Briefly write about discriminative method.	III	CO1
1D-12	How the internal structure of words can be modeled in connection with grammatical properties? Explain with examples. (Dec 2023)	II	CO1
1D-13	Describe topic boundary detection in a given document with suitable examples. (Dec 2023)	III	CO1

1D-14	What is finite state Morphology? Describe the approaches to design and implement morphological models. (June 2023)	I	C01
1D-15	Describe the procedure to find structure of documents along with its performance and complexity. (Jan 2024)	II	C01
UNIT-II			
Multiple Choice Questions		BT Level	Course Outcome
2C-1	What is the primary goal of syntax analysis in natural language processing (NLP)? A) To find synonyms of words B) To identify the structure and grammatical relationships between words in a sentence C) To translate text into another language D) To determine the sentiment of a sentence	II	C02
2C-2.	What is a treebank in the context of NLP? A) A large corpus of unlabeled sentences B) A structured database of syntactic structures annotated with grammatical information C) A model for semantic analysis D) A dataset for text classification	II	C02
2C-3.	Which of the following is a commonly used algorithm for syntactic parsing? A) K-means clustering B) CYK (Cocke-Younger-Kasami) algorithm C) Hidden Markov Models D) Naive Bayes classifier	I	C02
2C-4.	Which of the following best describes a "data-driven approach" to syntax analysis? A) Parsing based on predefined grammar rules B) Using statistical or machine learning models trained on annotated corpora to predict syntactic structure C) Handwriting recognition D) Syntactic analysis without any corpus-based information	V	C02
2C-5.	In a treebank, which of the following annotations are typically included? A) Part-of-speech tags B) Syntactic parse trees C) Constituency relations D) All of the above	III	C02
2C-6.	What is the key challenge addressed by ambiguity resolution in parsing? A) Determining the correct part-of-speech for each word B) Deciding between multiple valid parse tree structures for a sentence C) Identifying non-standard words D) Correctly tagging named entities	II	C02
2C-7.	Which of the following types of ambiguity does parsing in NLP aim to resolve? A) Word sense ambiguity B) Syntactic ambiguity (e.g., attachment ambiguity) C) Phonetic ambiguity D) Morphological ambiguity	VI	C02
2C-8.	Which of the following is a typical characteristic of a context-free	IV	C02

	grammar (CFG) used for syntactic parsing? A) It uses recursive rules to describe sentence structure B) It requires deep semantic analysis for disambiguation C) It is used to assign part-of-speech tags to words D) It is not suitable for ambiguous sentences		
2C-9.	Which of the following parsing strategies is typically used for efficient syntactic analysis? A) Top-down parsing B) Bottom-up parsing C) Chart parsing D) All of the above	V	C02
2C-10.	Which of the following is an example of an algorithm for parsing that handles ambiguity and considers all possible syntactic trees? A) Earley parser B) Shift-reduce parser C) Viterbi algorithm D) Naive Bayes classifier	III	C02
2C-11.	What is a significant challenge in parsing sentences across different languages? A) Different syntactic structures and word order B) The absence of part-of-speech tagging C) Lack of large corpora in some languages D) A and C	II	C02
2C-12.	In multilingual parsing, what is often required to handle different syntactic structures between languages? A) A unified lexicon for all languages B) Language-specific parsing models or treebanks C) A universal grammar D) A translation dictionary	I	C02
2C-13.	Which of the following is the correct representation of syntactic structure in a constituency grammar? A) Dependency tree B) Phrase structure tree C) Semantic representation D) Syllabic breakdown	III	C02
2C-14.	In dependency parsing, what is the primary focus of the analysis? A) The hierarchical structure of sentences B) The relations between words based on grammatical dependencies C) The segmentation of sentences into meaningful chunks D) The part-of-speech classification of each word	IV	C02
2C-15	Which of the following is a popular multilingual treebank project that aims to provide syntactic data across various languages? A) Universal Dependencies B) Penn Treebank C) Brown Corpus D) WordNet	V	C02
	Filling The Blank Questions		
2F-1	_____ refers to the process of analyzing the syntactic structure of a sentence, ensuring it follows the rules of grammar.	II	C02
2F-2	A _____ is a data structure that represents the syntactic structure of a sentence, usually as a tree.	II	C02
2F-3	_____ grammars are a type of formal grammar that provide rules for generating the structure of sentences in natural language.	I	C02
2F-4	In syntax analysis, the process of determining the hierarchical structure of a	I	C02

	sentence based on its grammatical components is called _____.				
2F-5	The _____ is the root of the parse tree, representing the entire sentence.		IV	C02	
2F-6	The set of all possible parse trees for a sentence is known as its _____.		I	C02	
2F-7	The _____ in a sentence often serve as the head of the phrase and play a key role in determining the structure.		V	C02	
2F-8	A _____ is a collection of syntactically annotated sentences that are used to train and evaluate parsing models.		VI	C02	
2F-9	Treebanks are important for _____ parsing, as they provide the data needed for statistical or machine learning models.		III	C02	
2F-10	In a data-driven approach to syntax, parsing models learn from _____ data rather than relying on handcrafted grammatical rules		I	C02	
2F-11	A _____ parsing algorithm works by trying to find the best tree that fits a set of syntactic rules and the sentence's input.		II	C02	
2F-12	_____ parsing involves building a parse tree from the leaves (words) to the root (sentence), considering possible substructures.		II	C02	
2F-13	_____ is a type of probabilistic model used in syntax analysis that helps resolve ambiguities by assigning probabilities to different parse trees.		IV	C02	
2F-14	_____ parsing models are often used in natural language processing to parse ambiguous sentences by considering the likelihood of different parses.		IV	C02	
2F-15	Multilingual parsing involves resolving _____ in different languages, as sentence structures vary greatly across linguistic systems.		II	C02	
	Match The Following Questions				
2M-1		List I	List II	II	C02
		1. Syntax Analysis	A. The process of analyzing sentence structure according to grammatical rules		
		2. Constituency Parsing	B. A tree structure that shows the grammatical structure of a sentence, with each node representing a constituent		
		3. Dependency Parsing	C. A parsing approach that represents the structure of a sentence as a tree with phrases as nodes		
		4. Syntactic Tree	D. A parsing approach that represents the relationships between words using directed links		
2M-2		List I	List II	I	C02
		1. Treebank	A. A treebank that annotates syntactic relationships between words based on dependencies		
		2. Syntactic Annotation	B. A corpus of sentences annotated with syntactic structures to be used for training parsers		
		3. Part-of-Speech Tagging	C. . The process of assigning syntactic categories such as noun or verb to words in a sentence		
		4. Dependency Treebank	D. A resource for labeling and storing syntactic structures of sentences in a formal format		

2M-3		List I	List II	II	CO2
		1. CYK Algorithm	A. A bottom-up parsing technique that uses a stack to reduce a sentence to its parse tree		
		2. Earley Parser	B. A parsing approach that uses a chart to store possible subtrees and parse all possible interpretations		
		3. Shift-Reduce Parsing	C. A top-down parsing algorithm that recursively breaks down sentences based on grammar rules		
		4. Chart Parsing	D. A parsing algorithm that builds a parse tree using dynamic programming for context-free grammars		
2M-4		List I	List II	I	CO2
		1. Syntactic Ambiguity	A. Ambiguity in the meaning of words, where a single word has multiple possible meanings		
		2. Lexical Ambiguity	B. The task of resolving multiple interpretations of words or phrases in context		
		3. Structural Ambiguity	C. Ambiguity in the sentence structure, where different syntactic structures are possible for a sentence		
		4. Disambiguation	D. Ambiguity that arises when a sentence can be parsed in multiple ways based on its structure		
2M-5		List I	List II	III	CO2
		1. Cross-Lingual Parsing	A. A parsing approach that works across languages without retraining for each language		
		2. Multilingual Treebanks	B. The process of automatically translating text from one language to another based on its syntactic structure		
		3. Universal Dependencies	C. The use of syntactic data for multiple languages to develop parsers that generalize across languages		
		4. Machine Translation	D. A syntactic framework designed to provide a consistent set of structures across multiple languages		

5 MARKS QUESTIONS

2D-1	Discuss multilingual issues on Natural Language Processing in detail. (Jan 2024)	II	CO2
2D-2	Explain the types of parsing.	II	CO2
2D-3	Explain Treebank's and its role in parsing. (June 2023)	II	CO2
2D-4	Elaborate the models for ambiguity resolution in parsing. (Jan 2024)	III	CO2

2D-5	Write a summary on types of parsing?	VI	CO2
2D-6	Differentiate between Constituency Grammar and Dependency Grammar.	IV	CO2
2D-7	Briefly explain various parsing algorithms.	V	CO2
2D-8	Write Shift-reduce parsing algorithm and trace it with an example. (Dec 2023)	I	CO2
2D-9	Explain syntactic analysis along with multiple examples.	II	CO2
2D-10	Differentiate between Shift-reduce parser and Chart parser.	IV	CO2
2D-11	Explain Chart Parser with an example.	II	CO2
2D-12	Explain syntax analysis using dependency graphs.(Dec 2023)	II	CO2
2D-13	With the help of neat diagram, Explain the representation of syntactic structure. (June 2023)	II	CO2
2D-14	What is Constituency Grammar?	II	CO2
2D-15	Briefly explain the types of derivations.	I	CO2

UNIT-III

Multiple Choice Questions		BT Level	Course Outcome
3C-1	What is the primary goal of semantic parsing in natural language processing (NLP)? A) To identify grammatical relationships between words B) To extract the meaning from a sentence and represent it in a formal structure C) To translate text into another language D) To determine the part-of-speech for each word	I	CO3
3C-2	What does "semantic interpretation" refer to in the context of NLP? A) Analyzing the syntax of a sentence B) Understanding the meaning of a sentence based on its structure and context C) Translating sentences from one language to another D) Categorizing words into their respective parts of speech	II	CO3
3C-3	Which of the following is a key feature of semantic parsing? A) Mapping linguistic inputs to a formal representation like a logical form or semantic graph B) Identifying the subject and object of a sentence C) Generating possible paraphrases for a sentence D) Extracting named entities from text	I	CO3
3C-4	What is the main challenge in semantic parsing? A) Decoding syntactic structure B) Mapping ambiguous linguistic structures to clear semantic representations C) Identifying part-of-speech tags D) Detecting negations in text	II	CO3
3C-5	Which of the following is an example of a formal representation used in semantic parsing? A) Parse tree B) Dependency graph C) Logical form D) Syllabic breakdown	I	CO3
3C-6	Which of the following semantic parsing paradigms focuses on extracting meaning through rules and logical structures? A) Statistical-based semantic parsing B) Rule-based semantic parsing	II	CO3

	C) Neural network-based semantic parsing D) Hybrid semantic parsing		
3C-7	In a data-driven approach to semantic parsing, which method is commonly used to learn semantic representations from large corpora? A) Deep learning B) Template-based parsing C) Heuristic-based systems D) Syntactic parsing alone	I	CO3
3C-8	Which of the following paradigms for semantic parsing aims to combine both rule-based and data-driven methods? A) Neural network-based parsing B) Hybrid semantic parsing C) Statistical parsing D) Syntactic parsing	II	CO3
3C-9	Which of the following is a common approach in semantic parsing using neural networks? A) Sequence-to-sequence models B) Bag-of-words models C) Statistical phrase structure grammar D) Markov chains	I	CO4
3C-10	What is the primary advantage of using a probabilistic or statistical approach in semantic parsing? A) Greater accuracy in word sense disambiguation B) Better handling of ambiguity in meaning and syntactic structures C) Easier interpretation of semantic structures D) Reduced computational complexity	III	CO4
3C-11	What does a word sense disambiguation (WSD) system aim to do? A) Detect and resolve ambiguity in the syntactic structure of a sentence B) Identify and assign the correct meaning to words based on context C) Translate words into other languages D) Parse sentences into logical forms	IV	CO4
3C-12	Which of the following is a well-known resource used for word sense disambiguation? A) WordNet B) BERT C) GPT-3 D) Penn Treebank	III	CO4
3C-13	In a word sense disambiguation system, which approach uses the surrounding context of a word to determine its meaning? A) Supervised learning B) Contextual disambiguation C) Morphological analysis D) Unsupervised learning	IV	CO4
3C-14	What is one key challenge in word sense disambiguation? A) Identifying the parts of speech of each word B) Correctly identifying word senses when there are multiple meanings for the same word C) Translating words from one language to another D) Disambiguating punctuation	III	CO4
3C-15	Which software framework is commonly used for training and deploying semantic parsing models based on deep learning? A) TensorFlow B) NLTK C) Spacy D) OpenNLP	V	CO4

Filling The Blank Questions					
3F-1	Semantic parsing involves mapping a sentence or utterance to a formal representation of its _____ meaning.			I	C03
3F-2	The goal of semantic interpretation is to determine the _____ of words and phrases in a given context.			I	C03
3F-3	Semantic parsing typically uses a combination of _____ models and rule-based methods to interpret language.			II	C03
3F-4	The _____ level of parsing aims to build representations that capture the meanings of sentences, such as logical forms or semantic networks.			II	C03
3F-5	In semantic parsing, the process of identifying which meanings of a word are applicable in a given context is known as _____.			I	C03
3F-6	A formal grammar-based semantic parsing system relies on pre-defined rules for mapping sentences to _____ representations.			II	C03
3F-7	In statistical systems, the goal is to learn semantic representations from large amounts of _____ data.			I	C03
3F-8	Neural network-based systems for semantic parsing have gained popularity due to their ability to automatically learn complex _____ from data.			II	C03
3F-9	One challenge in semantic parsing is to account for the _____ between syntax and semantics, ensuring that sentence structure aligns with its meaning.			I	C04
3F-10	A frame-based approach to semantic parsing involves mapping sentences to predefined _____ that capture roles and relationships.			III	C04
3F-11	Word sense disambiguation systems aim to determine the most appropriate _____ for a word based on context.			IV	C04
3F-12	A knowledge-based approach to word sense disambiguation uses _____ resources, such as WordNet, to identify word meanings.			III	C04
3F-13	Supervised learning methods for word sense disambiguation require large annotated corpora to train models for identifying word _____.			IV	C04
3F-14	One common issue in word sense systems is dealing with the _____ problem, where words can have multiple meanings in different contexts.			III	C04
3F-15	One popular software tool for semantic parsing is the Stanford Parser, which offers pre-trained models for both syntactic and _____ analysis.			VI	C04
	Match The Following Questions				
3M-1		1. Semantic Parsing	A. The analysis of sentences in terms of the truth conditions they satisfy	I	C03
		2. Compositional Semantics	B. The formal representation of meaning using logical systems or mathematical models		
		3. Truth-Conditional Semantics	C. A method of interpreting the meaning of sentences based on the meanings of their parts and how they combine		
		4. Formal Semantics	D. The process of converting natural language into a structured representation, such as logical forms or abstract syntax trees		

3M-2		1. Rule-Based Systems	A. Systems that use predefined rules and logic to interpret sentence meaning	II	C03
		2. Statistical Systems	B. Models that use deep learning techniques to automatically extract meaning representations from text		
		3. Neural Network Models	C. Systems that rely on training data and statistical methods to infer meanings from language		
		4. Hybrid Systems	D. Approaches combining different models to enhance accuracy and robustness in semantic interpretation		
3M-3		1. WordNet	A. A method for selecting the correct meaning of a word based on its context	I	C04
		2. Lesk Algorithm	B. A theory that suggests word meaning can be inferred from word co-occurrences in large corpora		
		3. Sense Disambiguation	C. An algorithm that determines word meaning by comparing definitions of possible senses		
		4. Distributional Semantics	D. A lexical database of English that organizes words into synonym sets (synsets)		
3M-4		1. OpenCyc	A. A framework for natural language processing that includes tools for semantic analysis, parsing, and named entity recognition	V	C04
		2. Stanford CoreNLP	B. A library for building deep learning models for natural language processing tasks, including semantic parsing		
		3. AllenNLP	C. An open-source tool that provides an ontology for common sense reasoning and semantic parsing		
		4. Spacy	D. A comprehensive NLP library for processing text that includes pre-trained models for semantic interpretation		
3M-5		1. Ambiguity	A. The issue of words or sentences having unclear or imprecise meanings that are hard to interpret	VI	C04
		2. Compositionality	B. The challenge of determining the correct meaning by considering how words combine within a sentence		
		3. Vagueness	C. The problem of a word having multiple meanings, which makes determining its exact sense difficult		
		4. Contextual Variation	D. The variation in meaning depending on the context in which a word or phrase is used		

3D-1	What are the types of representations of semantic parsing?	I	CO3
3D-2	What requirements are necessary for achieving semantic representation?	I	CO3
3D-3	What are the applications of semantic parser?	I	CO3
3D-4	Explain system architectures in semantic parsing.	II	CO3
3D-5	What is Word sense disambiguation? Explain different ways to identify correct sense of an ambiguous word.	I	CO3
3D-6	Explain working principle of Seq2Seq Neural Semantic Parsing with example. (June 2023)	II	CO3
3D-7	Describe word sense systems with details. (Jan 2024)	II	CO3
3D-8	Elaborate the Attention Mechanism for Neural Semantic parsing with neat diagram. (Jan 2024)	III	CO3
3D-9	Explain predicate and argument with examples.	II	CO4
3D-10	Write about the System paradigms.	VI	CO4
3D-11	Write short notes on Word sense in Semantic parsing.	III	CO4
3D-12	Explain with suitable example following relationships between word meanings: Homonymy, Polysemy, Synonymy, Antonymy, Hypernymy, Hyponymy	V	CO4
3D-13	Explain about various software research community for word sense?	II	CO4
3D-14	Briefly discuss various approaches for semantic parsing of natural language. (Dec 2023)	II	CO4
3D-15	a) Compare supervised and knowledge based systems for word sense disambiguation. (Dec 2023) b) Discuss the simplified Lesk algorithm. (Dec 2023)	III	CO4
3D-16	What are the types of representations of semantic parsing? Explain in details. (June 2023)	II	CO4

UNIT IV

Multiple Choice Questions			Course Outcome
4C-1	What does predicate-argument structure refer to in natural language processing? A) The relationships between words based on their syntactic roles B) The structure that describes the relationships between verbs (predicates) and their arguments (e.g., subject, object) C) The order of words in a sentence D) The mapping between words and their senses	I	CO5
4C-2	Which of the following is considered an argument in a predicate-argument structure? A) The verb "run" in the sentence "He runs." B) The word "run" as a noun C) The subject or object of a verb in a sentence D) The preposition in the sentence	II	CO5
4C-3	In a sentence like "Alice gave Bob the book," what is the predicate-argument structure? A) Alice → gave → Bob, the book B) Alice → gave → book C) Alice → book → gave D) Alice → gave → Bob → book	I	CO5
4C-4	Which of the following is the role of a predicate in a predicate-argument structure? A) To represent the subject of the sentence	II	CO5

	B) To represent the main action or event in the sentence C) To mark the boundary between clauses D) To link words together based on meaning		
4C-5	Which tool is often used to extract predicate-argument structures from text? A) Dependency parsing B) Named entity recognition C) Sentiment analysis D) Part-of-speech tagging	I	CO5
4C-6	What is the purpose of meaning representation systems in NLP? A) To determine the syntactic structure of a sentence B) To convert a natural language sentence into a formal representation that captures its meaning C) To generate possible paraphrases for a sentence D) To analyze the emotional tone of a sentence	II	CO5
4C-7	Which of the following is a commonly used meaning representation system? A) Universal Dependencies B) Conceptual Dependency Theory C) Hidden Markov Models D) Support Vector Machines	I	CO5
4C-8	In the Abstract Meaning Representation (AMR) framework, how is a sentence typically represented? A) As a syntactic tree B) As a sequence of words C) As a graph where nodes represent concepts and edges represent relationships between them D) As a list of named entities	II	CO5
4C-9	What is one advantage of using formal meaning representations like AMR or FrameNet? A) They allow for easy translation between languages B) They provide a rich representation of sentence meaning that is independent of syntax C) They simplify the task of part-of-speech tagging D) They only represent surface-level word meaning	I	CO5
4C-10	Which of the following best describes the FrameNet approach to meaning representation? A) It uses a set of predefined semantic frames to represent sentence meaning B) It builds a tree structure to represent grammatical relations C) It maps sentences to a predefined lexicon of word senses D) It uses an annotated corpus to train neural network models	III	CO5
4C-11	Which software tool is widely used for syntactic parsing, which is essential for extracting predicate-argument structures? A) TensorFlow B) SpaCy C) OpenNLP D) WordNet	IV	CO5

4C-12	Which of the following software tools is specifically designed to support the development of meaning representation systems like Abstract Meaning Representation (AMR)? A) NLTK B) AMR2.0 toolkit C) GATE D) AllenNLP	III	CO5
4C-13	Which of the following software libraries provides tools for working with semantic frames, such as those used in FrameNet? A) SpaCy B) NLTK C) Stanford NLP D) FrameNet Toolkit	IV	CO5
4C-14	What is the purpose of the PropBank corpus in relation to predicate-argument structures? A) It provides annotated data for training syntactic parsers B) It provides mappings between syntactic structures and their semantic roles (i.e., predicates and arguments) C) It represents sentences as a series of abstract meanings D) It provides a list of named entities	III	CO5
4C-15	Which of the following software packages is commonly used for semantic role labeling (SRL), a key task for understanding predicate-argument structure? A) SpaCy B) AllenNLP C) TensorFlow D) OpenNLP	V	CO5
Filling The Blank Questions			
4F-1	Predicate-argument structure refers to the relationship between a verb (the _____) and its associated noun phrases (the arguments).	I	CO5
4F-2	In a sentence, the subject and object are often the main _____ of a predicate, carrying out or receiving the action.	I	CO5
4F-3	The syntactic structure of a sentence often determines the _____ structure, where arguments are linked to specific roles like agent, theme, or patient.	II	CO5
4F-4	Arguments of a predicate can be _____ (like the subject or object) or _____ (like adverbs and prepositional phrases) based on the verb's requirements.	II	CO5
4F-5	In semantic analysis, the goal is to identify the roles that each argument plays, such as _____, patient, or experiencer.	I	CO5
4F-6	Meaning representation systems aim to transform sentences into formal representations, such as _____ or logical forms, to capture their meaning.	II	CO5
4F-7	A well-known meaning representation system that uses first-order logic to represent the meaning of natural language is called _____.	I	CO5
4F-8	Frame semantics represents meaning using _____, which capture the roles and relationships of entities in a particular context.	II	CO5
4F-9	Abstract meaning representation (AMR) is a meaning representation system that encodes the meaning of a sentence as a _____ graph.	I	CO5
4F-10	PropBank provides a semantic annotation framework for verbs, tagging the	III	CO5

	predicate-argument structure and offering detailed annotations for _____.												
4F-11	Natural language processing software often uses dependency parsers to identify the relationships between _____ in a sentence, which are crucial for constructing predicate-argument structures.	IV	CO5										
4F-12	The PropBank corpus is widely used in semantic role labeling , where software tools extract the semantic _____ of verbs from sentences.	III	CO5										
4F13	SpaCy is a popular open-source software library used for tasks like dependency parsing and named entity recognition in _____.	IV	CO5										
4F14	Semantic role labeling (SRL) systems typically rely on machine learning techniques to assign _____ roles to the arguments of predicates.	III	CO5										
4F15	Software tools like Stanford CoreNLP offer a range of NLP capabilities, including dependency parsing and semantic parsing for extracting _____ and meaning from text.	VI	CO5										
	Match The Following Questions												
4M-1	<table><tr><th>List I</th><th>List II</th></tr><tr><td>1. Predicate</td><td>A. A linguistic unit that expresses an action, state, or relation</td></tr><tr><td>2. Argument</td><td>B. A role assigned to an argument based on its semantic relationship with the predicate (e.g., agent, patient)</td></tr><tr><td>3. Thematic Role</td><td>C A syntactic element that fills the role of a participant in a sentence</td></tr><tr><td>4. Predicate-Argument Structure</td><td>D. . The set of relationships between a predicate and its arguments in a sentence</td></tr></table>	List I	List II	1. Predicate	A. A linguistic unit that expresses an action, state, or relation	2. Argument	B. A role assigned to an argument based on its semantic relationship with the predicate (e.g., agent, patient)	3. Thematic Role	C A syntactic element that fills the role of a participant in a sentence	4. Predicate-Argument Structure	D. . The set of relationships between a predicate and its arguments in a sentence	I	CO5
List I	List II												
1. Predicate	A. A linguistic unit that expresses an action, state, or relation												
2. Argument	B. A role assigned to an argument based on its semantic relationship with the predicate (e.g., agent, patient)												
3. Thematic Role	C A syntactic element that fills the role of a participant in a sentence												
4. Predicate-Argument Structure	D. . The set of relationships between a predicate and its arguments in a sentence												
4M-2	<table><tr><td>1. Logical Form (LF)</td><td>A. A semantic theory based on the assumption that word meanings are represented by conceptual structures (e.g., frames)</td></tr><tr><td>2. Frame Semantics</td><td>B. A syntactic structure where words are related based on dependencies rather than phrase structure</td></tr><tr><td>3. Dependency Grammar</td><td>C. A framework in which the meaning of a sentence is built up compositionally from the meanings of its parts</td></tr><tr><td>4. Compositional Semantics</td><td>D. . A formal representation of sentence meaning that captures truth conditions</td></tr></table>	1. Logical Form (LF)	A. A semantic theory based on the assumption that word meanings are represented by conceptual structures (e.g., frames)	2. Frame Semantics	B. A syntactic structure where words are related based on dependencies rather than phrase structure	3. Dependency Grammar	C. A framework in which the meaning of a sentence is built up compositionally from the meanings of its parts	4. Compositional Semantics	D. . A formal representation of sentence meaning that captures truth conditions	II	CO5		
1. Logical Form (LF)	A. A semantic theory based on the assumption that word meanings are represented by conceptual structures (e.g., frames)												
2. Frame Semantics	B. A syntactic structure where words are related based on dependencies rather than phrase structure												
3. Dependency Grammar	C. A framework in which the meaning of a sentence is built up compositionally from the meanings of its parts												
4. Compositional Semantics	D. . A formal representation of sentence meaning that captures truth conditions												

4M-3		1. OpenCyc	A. A resource that annotates sentences with frames, capturing roles and relationships within them	I	CO5
		2. FrameNet	B. A corpus providing a semantic role labeling system to mark arguments of predicates in a sentence		
		3. AllenNLP	C. A platform for building NLP models, including tools for meaning representation and argument structure parsing		
		4. PropBank	D. A project providing a formalized ontology for common-sense reasoning and knowledge representation		
4M-4		1. Agent	A. The argument that represents what is being affected or transferred in an event (e.g., the object in "She gave him the book")	V	CO5
		2. Patient	B. The argument that typically undergoes an action or change (e.g., the object in "kick the ball")		
		3. Experiencer	C. The argument that typically performs an action (e.g., the subject of a verb like "run")		
		4. Theme	D. The argument that experiences an event or state (e.g., the subject in "He enjoys the movie")		
4M-5		1. Ambiguity	A. The problem of determining the correct syntactic and semantic roles when a word or structure has multiple interpretations	VI	CO4
		2. Argument Identification	B. The challenge of transferring predicate-argument structures between languages with different syntactic structures		
		3. Role Labelling	C. The difficulty in correctly identifying and extracting the participants of a sentence		
		4. Cross-lingual Transfer	D. The task of associating a specific thematic role (like agent, patient) to each argument		
4D-1	Write a short notes predicate argument structure.			II	CO5
4D-2	What are the resources in Predicate argument structure?			II	CO5
4D-3	Write short notes on Frame net and Prop Bank.			I	CO5
4D-4	What thematic roles can be possible in Predicate and Argument Structure? Expalin. (Jan 2024)			II	CO5
4D-5	Elaborate the concept of Syntactic representation.			II	CO5
4D-6	What are the three tasks used to evaluate the system? Explain with an Algorithm.			II	CO5
4D-7	Make a comparison of FrameNet and PropBank For semantic role labeling. (Dec			I	CO5

	2023)		
4D-8	Differentiate between Phrase structure grammar and Combinatory Categorical Grammar.	IV	CO5
4D-9	List out the features used by the semantic chunker. Explain it.	IV	CO5
4D-10	Explain in detail about classification paradigms of semantic argument.	II	CO5
4D-11	Explain in detail about Meaning Representation in semantic parsing.	II	CO5
4D-12	How to overcoming the Parsing Errors in predicate argument structure.	III	CO5
4D-13	Describe the different types of representation systems in NLP with example. (Jan 2024)	II	CO5
4D-14	Appraise the significance of lambda calculus and illustrate its usage in meaning representation system. (Dec 2023)	III	CO5
4D-15	Describe the different types of predicates and Explain with suitable example to represent a predicate and argument. (June 2023)	III	CO5
UNIT V			
Multiple Choice Questions			Course Outcome
5C-1	What is discourse cohesion in natural language processing? A) The process of breaking a document into sentences B) The grammatical structure of individual sentences C) The links and relationships between sentences or parts of a text that contribute to its overall coherence D) The lexical similarity between words in a sentence	I	CO6
5C-2	Which of the following is an example of a cohesive device used in discourse? A) Synonymy B) Subject-verb agreement C) Pronouns and demonstratives that refer to previously mentioned entities D) Passive voice construction	II	CO6
5C-3	What does reference resolution in discourse processing refer to? A) Identifying the most relevant topic in a conversation B) Determining which word or phrase a pronoun or other referring expression (e.g., "he," "she," "it") refers to in a text C) Parsing the sentence structure D) Generating new discourse markers	I	CO6
5C-4	Which of the following is a challenge in discourse processing? A) Handling ambiguity in part-of-speech tagging B) Resolving ambiguities in word meanings C) Establishing coherent relationships between sentences and clauses D) Identifying the correct syntactic structure of a sentence	II	CO6
5C-5	What is the primary purpose of a language model in natural language processing? A) To parse syntactic structures of sentences B) To generate possible sentences or evaluate the likelihood of a given sentence C) To identify named entities in a sentence D) To determine the sentiment of a document	I	CO6
5C-6	Which of the following is the basic concept behind an N-gram language model? A) Modeling the probability of a word based on its surrounding context of N-words B) Using neural networks to generate sentences word by word	II	CO6

	C) Modeling the relationship between words and their parts of speech D) Using a fixed vocabulary to represent all words in a language		
5C-7	In an N-gram model, what does the "N" represent? A) The number of documents used in training the model B) The number of different words in a vocabulary C) The number of words in the context window used to predict the next word D) The number of language-specific rules applied during parsing	I	CO6
5C-8	Which of the following is a limitation of traditional N-gram models? A) They require a large amount of training data B) They are not capable of modeling context beyond a fixed window size C) They perform poorly in detecting part-of-speech tags D) They require high computational power to implement	II	CO6
5C-9	Which of the following is a common metric used to evaluate the performance of a language model? A) Precision B) F1-score C) Perplexity D) Word error rate	I	CO6
5C-10	What is "parameter estimation" in the context of language modeling? A) Assigning values to the parameters of a machine learning model B) Analyzing sentence structure to identify syntactic errors C) Measuring the complexity of the input text D) Assigning specific weights to words in a sentence based on their importance	III	CO6
5C-11	Which technique is commonly used for estimating the parameters (probabilities) in an N-gram language model? A) Maximum likelihood estimation (MLE) B) Deep learning C) Part-of-speech tagging D) K-means clustering	IV	CO6
5C-12	What is language model adaptation? A) Training a model on new data to improve its performance for a specific domain or task B) Preprocessing text data before feeding it into a model C) The process of improving the syntactic accuracy of a model D) Creating a model that generates random sentences	III	CO6
5C-13	Which type of language model uses neural networks to generate word sequences and predict the likelihood of future words in a sentence? A) N-gram model B) Markov model C) Recurrent Neural Network (RNN)-based model D) Hidden Markov Model	IV	CO6
5C-14	Which of the following is a challenge when building language models for languages with rich morphology, such as Finnish or Turkish? A) Handling large vocabulary sizes due to inflection and derivation B) Lack of available data for training C) Difficulty in assigning part-of-speech tags D) Lack of computational resources	III	CO6
5C-15	What is a key challenge in multilingual and cross-lingual language modeling? A) Differences in word order and syntactic structures between languages B) Identifying named entities in multiple languages C) Translating between languages using rule-based methods D) Handling polysemy in one language	V	CO6

Filling The Blank Questions													
5F-1	Discourse processing refers to understanding how sentences in a text are related and how their meaning extends beyond individual _____.	I	CO6										
5F-2	Cohesion in discourse is achieved through the use of linguistic devices like pronouns, conjunctions, and _____ that link different parts of the text.	I	CO6										
5F-3	Reference resolution involves determining what a pronoun or noun phrase refers to in a _____ context.	II	CO6										
5F-4	Anaphora is the linguistic phenomenon where a word (typically a pronoun) refers back to a _____ (usually a noun phrase) introduced earlier in the discourse.	II	CO6										
5F-5	Discourse cohesion is the relationship between sentences that ensures they are _____ together in terms of meaning and flow.	I	CO6										
5F-6	Language modeling is the task of estimating the likelihood of a sequence of _____ in a given language.	II	CO6										
5F-7	An N-gram model is a probabilistic language model that predicts the probability of a word based on the previous _____ words in the sequence.	I	CO6										
5F-8	The evaluation of a language model typically involves measuring metrics like _____ and perplexity to assess its predictive accuracy.	II	CO6										
5F-9	Parameter estimation in language models involves adjusting the model's parameters to best fit the data, often using techniques like _____ estimation.	I	CO6										
5F-10	Language model adaptation refers to the process of modifying an existing model to better suit specific _____ or contexts, such as a particular domain or topic.	III	CO6										
5F-11	The two primary types of language models are statistical models, such as N-grams, and neural models, such as _____ networks.	IV	CO6										
5F-12	Perplexity is a measure used to evaluate the effectiveness of a language model, where lower perplexity indicates a model with better _____.	III	CO6										
5F-13	Language-specific modeling problems arise when a language has unique grammatical features, such as _____ order or agglutination, which challenge standard modeling approaches.	IV	CO6										
5F-14	Multilingual language modeling involves creating a model that can handle multiple languages, often requiring adaptation of models to the specific _____ of each language.	III	CO6										
5F-15	Cross-lingual language modeling refers to building models that can work across different languages without requiring _____ for each language.	VI	CO6										
	Match The Following Questions												
5M-1	<table><tr><th>List I</th><th>List II</th></tr><tr><td>1. Cohesion</td><td>A. The overall organization of discourse into segments or units that contribute to meaning</td></tr><tr><td>2. Reference Resolution</td><td>B. The method of linking sentences or phrases to create a meaningful flow of ideas</td></tr><tr><td>3. Discourse Structure</td><td>C. The process of resolving relationships between linguistic elements, such as pronouns and their antecedents</td></tr><tr><td>4. Coherence</td><td>D. The logical flow or sense-making in a discourse, ensuring that the elements fit together cohesively</td></tr></table>	List I	List II	1. Cohesion	A. The overall organization of discourse into segments or units that contribute to meaning	2. Reference Resolution	B. The method of linking sentences or phrases to create a meaningful flow of ideas	3. Discourse Structure	C. The process of resolving relationships between linguistic elements, such as pronouns and their antecedents	4. Coherence	D. The logical flow or sense-making in a discourse, ensuring that the elements fit together cohesively	I	CO6
	List I	List II											
	1. Cohesion	A. The overall organization of discourse into segments or units that contribute to meaning											
	2. Reference Resolution	B. The method of linking sentences or phrases to create a meaningful flow of ideas											
	3. Discourse Structure	C. The process of resolving relationships between linguistic elements, such as pronouns and their antecedents											
4. Coherence	D. The logical flow or sense-making in a discourse, ensuring that the elements fit together cohesively												

5M-2	List I	List II	II	CO6
	1. N-Gram Model	A. The process of determining the most likely parameters (probabilities) for a given language model		
	2. Language Model Evaluation	B. The process of adjusting a language model to fit new or specific data sources		
	3. Parameter Estimation	C. A model that predicts the next word in a sequence based on the previous N words		
	4. Language Model Adaptation	D. A method used to assess the effectiveness of a language model in predicting language correctly		
5M-3	List I	List II	I	CO6
	1. Statistical Language Models	A. Methods used to handle zero or low-frequency events in language models		
	2. Neural Network-Based Models	B. Models that rely on probabilities calculated from data and are often used for speech recognition and text generation		
	3. Smoothing Techniques	C. A framework that applies Bayesian statistics to language model learning, where probabilities are updated based on prior knowledge		
	4. Bayesian Language Models	D. . Language models that utilize deep learning algorithms, such as LSTMs or Transformers, to predict word sequences		
5M-4	List I	List II	V	CO6
	1. Language-Specific Modeling Problems	A. The difficulty of developing models that can process multiple languages simultaneously		
	2. Cross-Lingual Language Modeling	B. The issue that arises when there is insufficient data to accurately estimate probabilities for certain words or contexts		
	3. OOV (Out of Vocabulary) Words	C. The challenge of handling words that do not appear in the training vocabulary		
	4. Data Sparsity	D. The problem of dealing with languages that have unique syntactic or morphological properties		
5M-5	List I	List II	VI	CO6
	1. Multilingual Language Modeling	A. The process of building models that work across many different languages by training on multilingual corpora		
	2. Transfer Learning	B. A scenario where a speaker alternates between two languages within a sentence or discourse		
	3. Cross-Lingual Embeddings	C. . A technique where a model trained in one language is applied to other languages with similar structures		
	4. Code-Switching	D. A type of embedding that represents words in multiple languages in a shared vector space		
5D-1	Explain three types of language models.		II	CO6
5D-2	Differentiate between Class based and variable length language models.		IV	CO6

5D-3	Differentiate between multilingual and cross lingual in the context of NLP. (June 2023).	IV	CO6
5D-4	Write a short note on Language-Specific Modeling Problem?	VI	CO6
5D-5	Explain coverage rate and perplexity?	II	CO6
5D-6	Write a short note on language modeling?	VI	CO6
5D-7	Differentiate between discriminative and syntax based language models.	IV	CO6
5D-8	Define Language Modeling . Explain N-Gram Model with its neat diagram. (Jun 2023)	IV	CO6
5D-9	Explain language modeling for morphologically rich languages.	II	CO6
5D-10	What is the need of using language modeling? Explain.	V	CO6
5D-11	Write a short note on n-gram notation?	III	CO6
5D-12	What is cohension in discourse analysis? Describe any four types of cohension in discourse. (Dec 2023)	V	CO6
5D-13	What is meant by language modeling? How does it work? Describe any two statistical language models. (Dec 2023)	III	CO6
5D-14	Explain three types of language models in details and it applications. (Jan 2024)	IV	CO6
5D-15	Elaborate discourse processing and its applications in details. (Jan 2024)	IV	CO6