



Shri Vile Parle Kelavani Mandal's

DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING

(Empowered Autonomous College Affiliated to the University of Mumbai)



Shri Vile Parle Kelavani Mandal's
Dwarkadas J. Sanghvi College of Engineering
(Autonomous College Affiliated to the University of Mumbai)

Scheme and detailed syllabus

Third Year B.Tech

in

Computer Science and Engineering

(IoT and Cyber Security with Block Chain Technology)

Honors Program in Data Analytics

(Semester V)

Prepared by:- Board of Studies in Computer Science & Engineering

(IoT and Cyber Security with Blockchain Technology)

With effect from the Academic Year: 2026-2027



Shri Vile Parle Kelavani Mandal's

DWARKADAS J.SANGHVI COLLEGE OF ENGINEERING

(Empowered Autonomous College Affiliated to the University of Mumbai)



Scheme for Third Year Undergraduate Program in Computer Science and Engineering (IoT and Cyber Security with Blockchain Technology) - Honors in Data Analytics : SEM V (Autonomous) Academic Year : 2026-27

Sr. no.		Course Code	Course	Teaching Scheme				Semester End Examination (SEE) - A						Continuous Assessment (SEE) - B					Aggregate (A+B)	Credits Earned		
				Theory (Hrs)	Practical (Hrs)	Tutorial (Hrs)	Credits	Duration (Hrs)	Theory	Oral	Practical	Oral & Practical	SEE Total (A)	Term Test 1 (TT1)	Term Test 2 (TT2)	Term Test 3 (TT3)	Term Total (TT1 + TT2 + TT3)	Term Work	CA Total (B)			
		SEM III																				
1		DJS23BH1201	Fundamentals of Data Mining	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	
		SEM IV																				
2		DJS23BH1251	Data Analytics and Visualization	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	
		DJS23BH1251L	Data Analytics and Visualization Laboratory	--	2	--	1	2	--	25	--	--	25	--	--	--	--	25	25	50	1	
		SEM V																				
3		DJS23BH1301	Natural Language Processing	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	
		DJS23BH1301L	Natural Language Processing Laboratory	--	2	--	1	2	--	25	--	--	25	--	--	--	--	25	25	50	1	
		SEM VI																				
4		DJS23BCH1351	Time Series and Forecasting Analytics	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	
		DJS23BH1351L	Time Series and Forecasting Analytics	--	2	--	1	2	--	25	--	--	25	--	--	--	--	25	25	50	1	
		SEM VI																				
5		DJS23BH1401	Business Intelligence	3	--	--	3	2	60	--	--	--	60	15	15	10	40	--	40	100	3	
		Total																				
				15	6	0	18	16	300	75	0	0	375	75	75	50	200	75	275	650	18	

Prepared by

Checked by

Dr. Vishwanath Jale

Head of the Department

Vice Principal

Principal



Program: B. Tech in Computer Science and Engineering (IoT and Cybersecurity with Block chain Technology) With Honors in Data Analytics									Third Year B.Tech		Semester: V	
Course: Natural Language Processing									Course Code: DJS23BH1301			
Course: Natural Language Processing Laboratory									Course Code: DJS23BH1301L			
Teaching Scheme (Hours / week)				Evaluation Scheme								
				Semester End Examination Marks (A)			Continuous Assessment Marks (B)				Total marks (A+ B)	
Lectures	Practical	Tutorial	Total Credits	Theory			Term Test 1	Term Test 2	Term Test 3	Total	100	
				60			15	15	10	40		
				Laboratory Examination			Term work					Total Term work
3	2	--	4	Oral	Practical	Oral & Practical	Laboratory Work	Tutorial / Mini project presentation Assignment		50		
				25	--	--		15	10		25	

Pre-requisite:

1. Probability and Statistical Inference
2. Python Programming.

Objectives:

1. Explain the fundamental concepts, evolution, and architecture of Natural Language Processing systems.
2. Apply text preprocessing and syntactic analysis techniques for processing natural language data.
3. Analyze semantic and discourse level structures to interpret meaning and context in natural language.
4. Evaluate and apply NLP techniques for solving real-world problems and emerging applications in IoT, Cybersecurity, and Blockchain.

Outcomes: On completion of the course, learners will be able to:

1. Explain the concepts, evolution, architecture, and challenges of Natural Language Processing systems.



2. Apply text processing techniques to preprocess textual data.
3. Analyze syntactic structures and sequence models for language modeling tasks.
4. Evaluate semantic analysis techniques for interpreting meaning in text.
5. Analyze discourse structures and coreference resolution techniques in Natural Language Processing.
6. Develop Natural Language Processing -based solutions for real-world applications.

Detailed Syllabus: (unit wise)		
Unit	Description	Duration
1	Introduction to Natural Language Processing (NLP): Definition and scope of NLP, History of NLP, Evolution of NLP (Rule-based → Statistical → Neural → LLMs), Traditional NLP vs LLM-based systems, Architecture / pipeline of NLP system, Levels of language processing, Ambiguity, challenges and limitations of NLP	7
2	Text Processing and Word Level Analysis: Words and tokens, Word types and word instances, Corpora and vocabulary, Tokenization techniques (rule-based, whitespace, etc.), Unicode and character encoding. Text normalization: Lowercasing, removing punctuation, Stop word removal. Regular expressions for text processing, Minimum edit distance (for spelling correction), N-grams: Unigram, Bigram, Trigram, N-gram language model (basic concept), Applications of N-grams: Text prediction, Spelling correction.	8
3	Syntactic Analysis and Sequence Modeling: Part-of-speech tagging and tag sets, rule-based POS tagging, stochastic POS tagging, issues in POS tagging (multiple tags and unknown words), context-free grammar, sequence labeling using Hidden Markov Models, Maximum Entropy models and Conditional Random Fields.	7
4	Semantic Analysis: Lexical semantics, word senses and semantic relations, homonymy, polysemy, synonymy and hyponymy, lexical resources such as WordNet, word sense disambiguation techniques, semantic role labeling and semantic representation in natural language processing.	7
5	Pragmatics and Discourse Processing: Pragmatics in natural language processing, discourse analysis, reference resolution, reference phenomena in discourse, syntactic and semantic constraints on coreference, coreference resolution techniques, applications of discourse processing in NLP.	6
6	Applications of NLP in Emerging Technologies: Text classification and sentiment analysis, named entity recognition applications, question answering systems, NLP applications in cybersecurity (phishing detection and threat intelligence analysis), NLP applications in IoT systems (voice commands and smart device interaction), NLP applications in blockchain analytics (cryptocurrency sentiment analysis and fraud detection), case studies of modern NLP systems.	7
Total		42



List of Laboratory Experiments:	
Sr. No.	Suggested Experiments
1	To Perform Text Preprocessing using NLTK
2	To Perform Stemming and Lemmatization:
3	To Perform Part-of-Speech Tagging
4	To Perform Chunking and Named Entity Recognition:
5	To Implement regular expressions for text processing.
6	To Implement N-gram Language Model
7	To Implement Minimum Edit Distance
8	To Implement TF-IDF Implementation
9	To Perform Word Sense Disambiguation
10	To Perform Sentiment Analysis
11	Mini Project (A mini project to be undertaken by a group of 4-5 students, focusing on topics related to the subject.)

Books Recommended:

Text Books:

1. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python, O'Reilly Media, 1st Edition, 2009.
2. Speech and Language Processing – Daniel Jurafsky & James H. Martin, 3rd Edition, Stanford University, 2023

Reference Books:

1. Christopher D. Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 1st Edition, 1999.
2. Jacob Eisenstein, Introduction to Natural Language Processing, MIT Press, 1st Edition, 2019.
3. Lewis Tunstall, Leandro von Werra and Thomas Wolf, Natural Language Processing with Transformers, O'Reilly Media, 1st Edition, 2022.

Online References:

1. Natural language toolkit- <https://www.nltk.org/>
2. Natural language Processing: <https://www.deeplearning.ai/resources/natural-language-processing/>
3. Speech and Language Processing book: https://web.stanford.edu/~jurafsky/slp3/?utm_source=chatgpt.com



Online Courses: NPTEL / Swayam:

1. Natural Language Processing by Prof. Pawan Goyal, IIT Kharagpur,
https://onlinecourses.nptel.ac.in/noc21_cs102/preview
2. Applied Natural Language Processing, By Prof. Ramaseshan R, CMI,
https://onlinecourses.nptel.ac.in/noc20_cs87/preview

Evaluation Scheme:

Semester End Examination (A):

Theory:

1. Question paper will be based on the entire syllabus summing up to 60 marks.
2. Total duration allotted for writing the paper is 2 hrs.

Laboratory:

Oral examination will be based on the entire syllabus including, the practical's performed during laboratory sessions.

Continuous Assessment (B):

Theory:

1. Term Test 1 (based on 40 % syllabus) of 15 marks for the duration of 45 min.
2. Term Test 2 (on next 40 % syllabus) of 15 marks for the duration of 45 min.
3. Term Test 3 Assignment / course project / group discussion / presentation / quiz/ any other for 10 marks.

Laboratory: (Term work)

1. Term Work shall consist of at least six practicals from the above list or any other practical that enhances students' conceptual understanding
2. The distribution of marks for term work shall be as follows:
 - i. Laboratory work (Performance of Experiments, Write-up): 15Marks
 - ii. Mini Project: 10 Marks

The final certification and acceptance of term work will be subject to satisfactory performance of laboratory work, and upon fulfilling minimum passing criteria in the term work.


Prepared by


Checked by


Head of the Department


Vice Principal


Principal