GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: BE-Minor/Hons.

Level: UG

Branch: Artificial Intelligence and Machine Learning

Course / Subject Code : BE06IAN011

Course / Subject Name : NLP using Python

w. e. f. Academic Year:	2024-2025
Semester:	6th
Category of the Course:	**Core Course**

Prerequisite:	• **Basic knowledge of Python programming** • **Fundamentals of linear algebra, probability, and statistics** • **Introductory understanding of machine learning concepts**
Rationale:	This course series provides foundational skills in generative AI models, natural language processing (NLP), and Python programming. Designed for developers, it offers progressive learning from basic to advanced concepts, enabling participants to build and optimize AI applications effectively for real-world challenges.

Course Outcome:

After Completion of the Course, Student will able to:

No	Course Outcomes	RBT Level
01	Understand the principles of generative and discriminative models and their applications in AI	U
02	Develop and fine-tune GANs and other generative models. Apply NLP tools and techniques to process and analyze text data effectively	A
03	Use Python libraries for NLP tasks, including tokenization, sentiment analysis, and language modeling.	A
04	Analyze potential challenges in GAN training and optimization.	N
05	Evaluate common issues in GAN and NLP model training and deployment.	E
06	Create AI models for practical applications to solve real-world problems using Python.	C

**Revised Bloom's Taxonomy (RBT)*

Teaching and Examination Scheme: (Not applicable, self-paced course)

Teaching Scheme (in Hours)			Total Credits L+T+ (PR/2)	Assessment Pattern and Marks				Total Marks
				Theory		Tutorial / Practical		
L	T	PR	C	ESE (E)	PA / CA (M)	PA/CA (I)	ESE (V)	
5	0	0	5	100	0	0	0	100

Course0 Content:

Unit No.	Content	No. of Hours	% of Weightage
1.	**Generative models for developers,** Introduction to Generative models, Generative model, Applications of Generative model, Approaches of Generative model, Generative Adversarial Network -GAN, Naive Bayes generative model, Generative Adversarial Network- GAN architecture, Training the GAN, Applications of GANs, Types of GANs, Deep Convolutional GAN, Steps for building DCGAN, Additional GAN architectures, Additional GAN Architecture, Conditional GAN Architecture, AuxilIANy classifier GAN, Wasserstein GAN, Challenges and performance issues in GAN models, Challenges in Training GAN, Performance evaluation for GAN	19	20
2.	**Natural Language Processing for developers -** Basic NLP Concepts, Why NLP, NLP Pipeline, Understanding Textual Data, Corpus, Wordnet, Wordlist, Lexical resources, POS Tagging, Annotation, NER, Bag of Word and TF_IDF, Working with multiple documents, Distributions of N-grams, Lookup tagger, Tagging pipeline and backoff, Basic NLP Applications using Machine Learning, SMS Spam Classifier using Decision Trees and Naive Bayes, Elementary Sentiment analysis, Elementary Chatbot using basic NLP, Elementary Topic Modeling, Applications of NLP, Advanced NLP Concepts, Why advanced NLP, Embedding Words, Word2vec, How to Train Word2Vec on Text data, Visualization of Trained model, Word Vector, Sequential Modeling for NLP using RNN, What is RNN ? How Business can benefit from RNN? Language Modeling using RNN, Understanding RNN with Text Generation/ word prediction/stock Market prediction example, Challenges / Limitation of RNN for NLP problems, Recurrent Neural Network, Sequential Modeling for NLP using LSTM, How LSTM overcome Vanishing gradient problem, Training using LSTM for Language Modeling - stock Market prediction/ Word prediction, Tuning the parameters in LSTM, Bi-directional LSTM, Comparison of RNN vs LSTM, GRU, Benchmarking for various solutions, Characteristics of NLTK, SPACY, core NLP and SPARK NLP, Need For Evaluation Metric, Benchmarking for Preprocessing of Text, Benchmarking Solutions, Topic Modeling, What is topic analysis?, Transformers, BERT, Conversational Interface and Chatbot, Conversational Chat Bot, **Guided Project**	14	16
3.	**Natural Language Processing using Python** - Why NLP? Understanding Textual Data, Elements of Text, , Distributions and n-grams, Operations on Text, - Understanding textual data, Annotation, Text annotation using chunking and named entity recognition, Annotation, Exercise - Annotations, Tagging, POS tagger, Default tagger, Lookup taggers - Unigram and N-gram	16	18

	taggers, Tagging pipeline and backoff, Working with multiple documents, Corpus, Bag of words and TF-IDF models, - Working with multiple documents, Lexical resources, Wordlists, WordNet, NLP and Machine Learning - Applications, NLP Applications, NLP Pipeline, SMS Spam Classifier using Decision Trees and Naive Bayes, Sentiment analysis, Chatbot, Topic Modeling, Entity Relationship Extraction, Conditional Random Fields, Applications of NLP, **Capstone Project**		
4.	**Programming Fundamentals using Python - Part 1 -** Variables, Datatypes & Operators in Python, what is a programming language? Datatypes, Variables, Variables & Datatypes, Operators, Implicit & Explicit Type Conversion, Variables & Datatypes, Coding Standards & Formatting Output, Formatting Output Activity	13	15
5.	Introduction to Functions, Basics of Functions, Basics of Functions Exercise, Flow of Execution in Functions, 'return' in Function, Function Invocation, Selection Control Structures, Nested Selection Control Structure, Iteration Control Structures Syntax, 'while', 'for' with 'in', Iteration Control Structures 'for' with 'range', 'for' with 'range', Iteration Control Structures Nested Loops, Nested Loops, Nested Loops of Different Types, Iteration Control Structures 'break' & 'continue', 'break' & 'continue', Iteration Control Structures Testing Loops Quiz, Exercise on Iteration in Python, Programming Fundamentals, Eclipse Plugin, Introduction to Eclipse Plugin, Configuring Eclipse Plugin, Eclipse Input Tip to Handle Infinite Loop, Collections and List in Python, List, List Iteration, List Searching, List Slice, List of Lists, List Functions,	14	16
6	Tuple in Python, Tuple, String in Python, String, String Functions, Choosing Between List, Tuple and String, Debugging, Eclipse Debugger, Programming Fundamentals: Dictionary in Python, Dictionary, Set in Python, Set, Libraries & Functions in Python, Random Library, Math Library, Date & Time Functions	**13**	**15**
	Total	**89**	**100**

Suggested Specification Table with Marks (Theory): Given here tentative, which may vary as per Author and Course.

Distribution of Theory Marks (in %)					
R Level	**U Level**	**A Level**	**N Level**	**E Level**	**C Level**
10%	**15%**	**35%**	**20%**	**10%**	**10%**

Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)

* * * * * * *