BCSE0704-COMPUTATIONAL LINGUISTICS AND NATURAL LANGUAGE PROCESSING

Objective: *The course objectives for each of the*

Credits: 04 **Semester V** **L-T-P: 2-0-0**

Module No.	Content	Hours
I	**Introduction to NLP:** linguistic terminology- Morpheme, Grapheme, Phoneme, Classical Approaches of NLP, Understanding linguistics – Morphology, Syntax, Semantics, Pragmatics. Basics of Text processing, principles of text analysis: Regular Expression, word tokenization, Word normalization- Lemmatization and stemming, Stop words and Key words identification, Introduction to N-Gram models, Bag of Words Representation, PoS tagging, **Statistical Approaches:** Statistical parsing, Approaches to parsing, Words & Vectors-Word 2 Vec concepts, TF-IDF computation, Inverted Index construction, Document Incidence Matrix construction, Text similarity methods—Similarity coefficient, Jaccard similarity, Cosine similarity.	20
II	**Text Classification:** Spam detection, Language Identification, Sentiment classification. Classification Methods- supervised and unsupervised. Machine learning in action: document classification information extraction- Named Entity Recognition. **Applications of NLP:** Information retrieval in NLP, Design Feature of IR systems, Question Answering system., QA System types, Opinion Mining, Sentiment analysis, Recommendation system, Machine Translation, Word sense disambiguation, Word embedding concept, Duplicate detection, Concept of Shingling, Ontology construction and classification, Performance and correctness measures- Precision, Recall and f-Measure. Emerging Applications of Natural Language Generation in Information Visualization, Education, and Healthcare.	20

Text Books:

- The handbook of Computational Linguistics and Natural Language Processing, " Alexandar Clark, Chris Fox and Shallom Lappin, Wiley.
- Handbook of NLP , Nitin Indurkhya, Fred J. Damerau, CRC Press

Reference Books:

- NLP with Python, Steven Bird, Ewan Klein, and Edward Loper, Oreilly
- Speech and Language Processing, Jurafsky, Daniel and James Martin, Prentice-Hall,

Head of the Department
Computer Engineering & Applica
Institute of Engineering & Techn
GLA University, Mathura