APPLIED NATURAL LANGUAGE PROCESSING

PROF. RAMASESHAN R
Department of Computer Science and Engineering
Chennai Mathematical Institute

TYPE OF COURSE	: Rerun | Elective | UG/PG
COURSE DURATION	: 12 weeks (20 Jul'20 - 9 Oct'20)
EXAM DATE	: 17 Oct 2020

PRE-REQUISITES : Essential – Algorithms, Python Proficiency, Elementary probability and statistics, Linear Algebra, basic understanding of machine learning

COURSE OUTLINE :
A major portion of communication now is through text and any organization has more than 90% of its content in the unstructured form. Natural Language Processing (NLP), an important part in Artificial Intelligence, is one of the important technologies that would help in activities such as classification, retrieving and extraction of information, identifying important documents, etc. Students will gather knowledge in the fundamentals of NLP, methods and techniques and gain skills to use them in practical situations.

ABOUT INSTRUCTOR :
Prof. Rameseshan R is currently working as a Visiting faculty at Chennai Mathematical Institute and teaches this course to the students there. He has more than 30 years of experience in Research and Development, Teaching, Product Development, Information Technology, Innovation and Convergence.

COURSE PLAN :

Week 1**:** Introduction to language processing – tokens, sentences, paragraphs

Week 2**:** Regular expressions - extraction of information using Regex

Week 3**:** Document Similarity measures - Cosine and cluster measures

Week 4**:** Spelling correction - Edit distance

Week 5**:** Information retrieval, extraction

Week 6**:** Document Classification, Clustering, topic modeling techniques

Week 7**:** Vector Space Model - word vectors, GloVe/Word2Vec model, word embedding

Week 8**:** Text Classification, Clustering, and Summarization

Week 9**:** Machine Learning, Perceptron

Week 10**:** Back Propagation, Recurrent Neural network relevant to NLP

Week 11**:** Machine Translation, Language Generation

Week 12**:** Applications – Sentiment Analysis, Spam Detection, Resume Mining, AInstein