

A Certificate Course in “Computational Linguistics” [for TWO credits]

(15th April to 15th June 2021)

Syllabus

Unit 1 Introduction to Computational Linguistics

Computational Linguistics, Artificial Intelligence, A short history of Computational Linguistics is provided with traditional grammar and Computer Interface on Language tools. The evolution of Natural Language based Software development and Cognitive Science.

Unit 2: An Introduction to Text Analysis

- Text and Discourse definition, characteristics, various uses and different types of corpora to be used in the computer processing of Natural Language Texts
- The methods of corpus building especially machine readable corpus. Different corpora in use like Brown Corpus/LOB Corpus and Bank of English.
- Introduction to Indian Language Text Processing Scopes and Works
- Introduction to various tools of Text Analysis like WordNet, Spell Checker, Parts-of-Speech (POS) Tagging, Parser, Morphological Analyzer, Stemmer etc.
- Introduction to Linguistic standards and resources for Language Technology

Unit 3: Major Applications of Language Technology: Machine Translation and Information Retrieval

Machine Translation, ambiguity in natural languages with examples, different approaches (rule-based approach, statistical and data-driven approach).

Unit 4: Major Applications of Language Technology: Speech Technology

Speech synthesis, Speech recognition

Unit 5: Practical Application Demonstration of Language Technology: Various

Computational Tools Demonstration and Hand Holding Session on all Language Technology topics of Unit 2-4.