**Natural Language Processing (NLP)**

**Session 1**

1. **NLP**: Concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. NLP combines computational linguistics—rule-based modeling of human language—with statistical, machine learning, and deep learning models. Together, these technologies enable computers to process human language in the form of text or voice data and to ‘understand’ its full meaning, complete with the speaker or writer’s intent and sentiment.  
     
   Applications include managing big data i.e., Google Search and analysing social media content, and dialogue systems/chatbots (spoken and typed interfaces).

Read up on Google Translate statistics.

1. **Managing Big Data**:
   1. *Classification*: identify relevant content / quickly assess this content.
   2. *Extract*: structured information from unstructured textual data.
   3. *Summarize*: compressing the full text into smaller readable summaries either from a single source (single-document summarization) or from a collection of articles (multi-document summarization). Main approaches:
      1. *Extractive*: Select key sentences/phrases for summary.
      2. *Abstractive*: Re-generate a summary based on the meaning of the text.
2. **Sentiment Analysis**: