Natural Language Processing(NLP)

Outline

➢ **Use of NLP**

➢ **Introduction**

➢ **Why NLP ?**

➢ **Areas in NLP**

➢ **Future of NLP**

➢ **References**

Use of NLP

➤ **Have you ever used any NLP products/ NLP Powered tools ?**
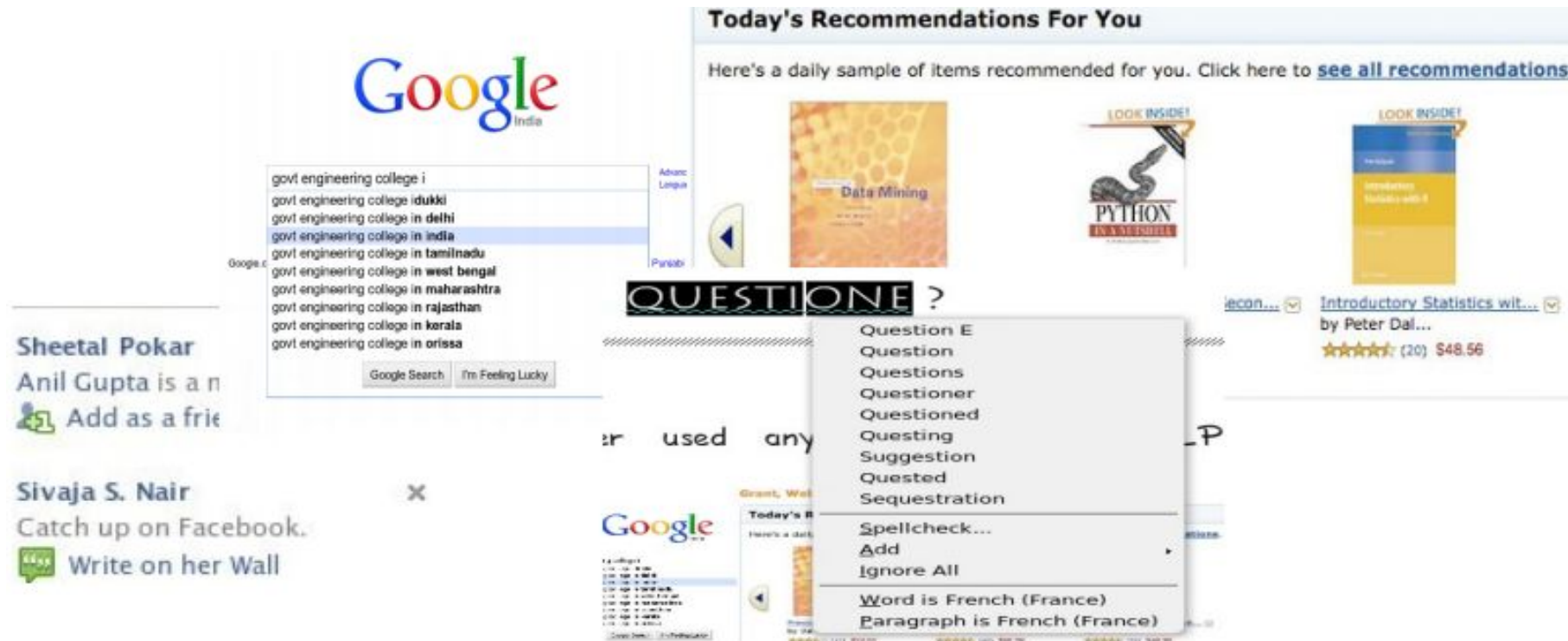
Natural Language Processing

➢ **A sub-field of Artificial Intelligence (AI)**

➢ **An inter disciplinary subject.**

➢ **Aim:**

➢ **To build intelligent computers that can interact with human being like a human being !!**

Definition

Natural language processing (NLP) is a field of computer science, artificial intelligence (also called machine learning), and linguistics concerned with the interactions between computers and human (natural) languages. Specifically, the process of a computer extracting meaningful information from natural language input and/or producing natural language output.

Why NLP

➢ **Huge amounts of data**

  Internet => at least 20 billions pages

  Text data => web sites, blog, tweets ....

  Audio data => speech ....

➢ **Impossible to read all data and then extract information**

➢ **Applications for processing large amounts of texts require NLP expertise.**

Example

➢ **News: AN EARTHQUAKE struck Indonesia today - a strapping 7.7 magnitude earthquake that struck early today off the northern coast of the island of Sumatra. It caused minor damage and there are no reports of any deaths, although electricity was interrupted in several places.**

➢ **Information From the News:**

Location : Indonesia
Magnitude: 7.7
Region: Sumatra (Northern Cost)
Deaths: Nil
Damage: Minor

Major Areas of Research & Development

➤ **Text Processing**
         - processing of raw text.

➤ **Morphological Analysis – study of word formation (how words are built up from smaller pieces)**
         - Washing = Wash + ing
         - Running = Run + ing

➤ **POS Tagging**
         >> text = word_tokenize("today is a beautiful day")
         >> nltk.pos_tag(text)

Output: [('today', 'NN'), ('is', 'PRP'), ('a', 'DT'), ('beautiful', 'JJ'), ('day', 'NN')]

Major Areas of Research & Development

➢ **Machine Translation**

    Translating content in one natural language to another natural language. Example :

    তুষার খুব ভাল ছেলে **=> Tushar is a very good boy.**

➢ **Parsing**

    Identifying sentence structure
    S ☐ NP + VP

➢ **Text to speech**

    Converting electronic text to digital speech.

Major Areas of Research & Development

➢ **Automatic Speech Recognition**

   Automatic transcription of spoken content to electronic text

➢ **Speech to speech translation**

   Translating spoken content from one language to another in real time or offline.

industrial Applications

➢ **Search Engines**

➢ **Fraud Detection**

➢ **Information Extraction**

➢ **Collaborative filtering**

➢ **Computational Advertising**

➢ **Advanced Text Editors**

➢ **Sentiment Analysis**

➢ **Opinion Mining**

Future of NLP

- ➢ **Semantic Web/Search**
- ➢ **Advanced Sentiment Analysis / Opinion Mining**
- ➢ **Machine Translation**
- ➢ **Advanced Speech Processing Application**
- ➢ **Social Network Analysis**
- ➢ **Collective Intelligence**

NLP in other Domains

➢ **Bio-Medical**

➢ **Forensic Science**

➢ **Advertisement**

➢ **Politics**

➢ **E-governance**

➢ **Business Development**

➢ **Marketing**

➢ **and where ever we use language !!!**

References

- "Speech and Language Processing"

    by Daniel Jurafsky & James H. Martin

Thank You ☺