FABER-CASTELL

Date

Page No.

Natural Language Processing.

Module → 1, 2, 3, 4.1

Mod 1 → Intro → - History
System
Levels
Knowledge in LP
Ambiguity.
Stages
Challenges
Applications.

—x—

Goals of NLP

- Deep understanding of broad language construct.
- Achieve human like comprehension of texts/langs.
- System to understand, draw interfaces from
  - summarize
  - translate & generate accurate and natural human text and language
  - Opinion dynamics
  - Spam detection
  - Text summarization.

An area of computer science that deals with methods to analyse model, and understand human language.

Core applications: 1.) Email platforms
2.) Voice based assistants
3.) Modern search engines.
4.) Machine translation services

Date
Page No.

NLP Tasks

1.) Language Modeling
- predict what the next word will be.
- learn the probability of a sequence of words appearing in a given language.
- speech recognition, machine translation, spelling correction.

2.) Text classification
- bucketing the text into known set of categories based on its content.
- email spam identification, sentiment analyses.

3.) Info. extraction
- extracting relevent info. from text.
- calender events from e-mail.

4.) Info. retsiral.
- finding documents relevent to user query from a large collection.
- google search. / scholar.

5.) conversational agent
- building dialogue systems that can converse in human languages.
- Alexa, Siri, etc.

6.) Text Summarization :
- Create short summaries of larges documents.

Date
Page No.

7.) Question answering :
- Build a system that can automatically answer questions based in Natural language.

8.) Machine Translation :
- Translate one language text to another.
- Google translate.

9.) Topic Modeling :
A topic model is one that automatically discovers topics occuring in a collection of documents.

→ Building blocks of language.

1.) Phonemes :
Smallest unit of sound in a language

2.) Morphemes :
- Smallest unit of language that has a meaning.
- Combination of phonemes.

3.) Lexemes :
Structural variations of morphemes related to one another by meaning.

4.) Syntax :
Set of rules to construct grammatically correct sentences out of words and phrases in a language.

5.) Context :
How various parts in a language come together to convey a particular meaning.

Date
Page No.

- Semantics:
  Direct meaning of the words & sentences without external context.
- Pragmatics:
  Adds world knowledge & external context of the conversation to enable us to infer implied meaning.

Complex NLP → • Sarcasm Detection
↓ • Summarization

use Context

Why is NLP challenging?

→ • Ambiguity
• Creativity of human language.

Components of NLP

- NLUnderstanding (NLU)
- NLGeneration (NLG)

1.) NLU:
Mapping the given input in NL into useful representation.

2.) NLG:
producing meaningful phrases & sentences.
Involves ① Text planning.
② Sentence planning.
③ Text realization.

FABER-CASTELL
Date
Page No.

NLU is harder than NLG

Ambiguity Types

1. Lexical
2. Syntax
3. Referential

Stages in NLP

Lexical (analyzing the structure of words)
↓
Syntactic (grammer & assanging words to make sense)
↓
Semantic (check for meaningfulness)
↓
Disclosure (meaning of any sentence depends upon meaning of sentence before it)
↓
Pragmatic (reinterpreted on ~~words~~ what it actually meant, real world knowledge)