

# **PERTEMUAN 10**

## **NATURAL LANGUAGE PROCESSING (PEMROSES BAHASA ALAMI)**

# Definisi

Pemroses bahasa alami (Natural Language Processing/NLP) adalah suatu aplikasi (program) dalam bidang AI yang dapat mengartikan suatu bahasa baik bahasa tulisan maupun bahasa lisan atau memproses masukan yang berupa bahasa menjadi suatu informasi atau pengetahuan

Yang menjadi pembahasan bukan bagaimana bahasa diinput atau dimasukkan kedalam program, tetapi lebih kepada bagaimana mengartikan suatu bahasa atau mengcopy / mengambil informasi/pengetahuan dari suatu bahasa.

# Pembagian NLP

Masalah pemrosesan bahasa alami dibagi menjadi dua bagian besar, yaitu :

1. Pemrosesan Naskah Tertulis  
menggunakan pengetahuan tentang leksikal, syntax, dan semantik
2. Pemrosesan Bahasa Lisan  
menggunakan semua pengetahuan dari pemrosesan naskah tertulis ditambah pengetahuan tentang phonology.

# Masalah dalam NLP

Beberapa masalah yang dihadapi dalam pemrosesan bahasa alami antara lain adalah :

1. Suatu kalimat sering kali tidak lengkap, artinya tidak memberi informasi yang jelas atau lengkap
2. Satu kalimat dapat memiliki lebih dari satu pengertian dalam konteks yang berbeda
3. Tidak ada program pemroses bahasa alami yang cukup lengkap karena bahasa selalu berkembang, kosa kata selalu bertambah.
4. Bisa terdapat lebih dari satu cara (lebih dari satu kalimat) untuk mengungkapkan hal(maksud) yang sama.

# Tahapan Proses

Untuk memproses bahasa alami diperlukan 5 langkah sebagai berikut :

## 1. Analisis Morpology

Pada tahap ini dilakukan analisa untuk setiap kata dan komponen yang dimiliki tiap kata termasuk token non kata seperti spasi, tanda baca, tanda pemisah.

## 2. Analisis Sintax

Pada tahap ini sederetan kata disusun kedalam struktur yang memperlihatkan bagaimana hubungan satu kata dengan kata lainnya. Deretan kata akan ditolak bila tidak memenuhi aturan penyusunan kata yang ada

### 3. Analisis semantik

Pada tahap ini struktur deretan kata yang sudah terbentuk akan diberi arti. Dengan kata lain pemetaan dibuat antara struktur sintax dengan object yang berhubungan.

- **Penyatuan Arah (konteks)**

Pada tahap ini arti dari suatu kalimat disesuaikan dengan kalimat-kalimat lain, karena arti dari suatu kalimat biasanya berhubungan dengan kalimat sebelumnya dan kalimat sesudahnya.

### 5. Analisis Pragmatis

Struktur yang terbentuk menghasilkan interpretasi ulang dari apa yang sudah dikatakan atau ditulis sebelumnya dengan arti yang sebenarnya.

# Grammar dan Parsers

- Grammar adalah suatu aturan yang menentukan bagaimana suatu kalimat dalam suatu bahasa dibentuk. Grammar berisi kumpulan syntax yang baku/benar dari suatu bahasa.

Contoh : Dalam bahasa Indonesia, suatu kalimat biasanya terdiri dari

Subject-Predikat-Object-Keterangan

- Parsers adalah suatu metode atau suatu program (sering disebut suatu mesin) yang dapat memproduksi/menghasilkan kalimat atau bahasa yang sesuai dengan Grammar yang sudah ditentukan atau diinginkan. Parsers juga dapat memeriksa apakah suatu kalimat yang dimasukkan sesuai dengan Grammar atau tidak.

# Jenis Parsers

Parsers terdiri dari dua jenis, yaitu :

1. Top-Down Parsing

memulai proses parsing dari simbol start dan menggunakan aturan grammar sampai simbol-simbol terminal pada tree terhubung ke komponen kalimat yang di parsing

2. Bottom-Up Parsing

memulai proses parsing dari kalimat yang akan di parsing dan menggunakan aturan grammar secara terbalik untuk memproduksi kata menjadi terminal, terminal menjadi kalimat sampai tree/ pohon lengkap dan simbol start tercapai



# THE END