

**PROPOSAL PROYEK AKHIR MATA KULIAH
PEMROSESAN BAHASA ALAMI**



Indonesian Named-entity Recognition

OLEH:

12S17028	Andri Reimondo Tamba
12S18006	Mei Kristina Panjaitan
12S18013	Yudika Purba
12S18021	Lastri Sari Naomi Marbun

**PROGRAM STUDI SARJANA SISTEM INFORMASI
INSTITUT TEKNOLOGI DEL NOVEMBER 2021**

DAFTAR ISI

DAFTAR GAMBAR.....	1
DAFTAR TABEL	2
BAB 1 PENDAHULUAN	3
1.1 Latar Belakang.....	3
1.2 Tujuan Penelitian	3
1.3 Manfaat Penelitian.....	4
1.4 Ruang Lingkup.....	4
BAB II ISI	5
2.1 A.....	5
2.2 B	5
2.3 C	5
2.4 D.....	5
DAFTAR PUSTAKA	2

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1 PENDAHULUAN

Bagian ini berisi latar belakang, tujuan, manfaat, ruang lingkup. Tim pengembang, serta kepentingan utama dalam pengerjaan Proyek dalam mata kuliah PBA (Pemrosesan Bahasa Alami) dalam topik Indonesian Named-entity Recognition

1.1 Latar Belakang

Named-entity Recognition merupakan bagian dari riset Natural Language Processing (NLP) yang digunakan untuk mengekstrak informasi serta nama orang, organisasi, lokasi dan waktu. Dengan tujuan misalkan untuk memperoleh informasi dalam dokumen teks secara manual, manusia harus membaca seluruh isi dokumen, apabila dokumen teks memiliki teks panjang diperlukannya waktu sangat lama dalam mendapatkan informasi tersebut dengan itu Named- entity Recognition ini memperoleh informasi dengan waktu yang singkat. [1]

Dalam pendekatan Named-entity Recognition dalam proyek ini menggunakan entity tagging namun ada kekurangan yang dimiliki yaitu pada proses automatic tagging dimana tidak dapat melakukan tag entitas pada kata serta frasa yang dimiliki entitas, untuk meningkatkan kemampuan automatic tagging dengan melakukan POS-Tagging dengan menambahkan beberapa aturan pada automatic tagging. Dengan Post-Tagging ini akan memperoleh seluruh kata atau frasa yang ketika akan melakukan pengecekan dalam rule yang disediakan pada tag entitas berdasarkan aturan yang digunakan.

1.2 Tujuan Penelitian

Tujuan penelitian ini adalah untuk menghasilkan model Named entity Recognition Bahasa Indonesia yang dapat digunakan untuk mempermudah proses dalam menentukan jenis entitas pada kata atau frasa yang terdapat dalam kalimat atau teks Bahasa Indonesia.

Dalam proyek ini adanya manfaat penelitian yang diperoleh adalah sebagai tahapan awal information Extraction bahasa Indonesia.

1.3 Manfaat Penelitian

Adapun yang menjadi manfaat proyek adalah sebagai berikut:

a. Penulis dapat menambah wawasan dan pengalaman secara langsung dalam

membangun sebuah model untuk untuk mengekstrak informasi serta nama orang, organisasi, lokasi dan waktu.

b. Sebagai referensi, jika ingin dilakukan pengembangan terhadap proyek.

1.4 Ruang Lingkup

Ruang lingkup dari proyek ini dibatasi dengan data awal yaitu Singgalang dataset bahasa indonesia. Data set ini di hasilkan menggunakan DBpedia modifikasi dataset DEE (Alfina et al., 2017), dimana dataset ini akan digunakan sebagai data untuk menghasilkan Named Entity Recognition (mengklasifikasi entitas) dalam teks dengan menggunakan Conditional Random Field dan Pos Tagging pada sistem yang akan dibangun.

BAB II ISI

2.1 A

2.2 B

2.3 C

2.4 D

DAFTAR PUSTAKA

- [1] *NAMED ENTITY RECOGNITION (NER) BAHASA INDONESIA MENGGUNAKAN CONDITIONAL RANDOM FIELD DAN POS-TAGGING, 2018*
- [2] ---
- [3] ---
- [4] --
- [5] --
- [6] --

