**NAMED ENTITY RECOGNITION**

BAHASA INDONESIA

Dalam penggunaan pada anago, format data training harus dirubah menjadi format BIO (seperti yang tertera pada data contoh). penggunaan awalan label “B-” dan “I-” seperti B-LOC, I-LOC, B-PER, I-PER, B-QTY, I-QTY dan B-TIME, I-TIME wajib di anago karena di codenya mendeteksi “B-” dan “I-“. Sedangkan untuk bagian belakangnya bebas. Sehingga kita dapat merubah atau menambah entity sesuai dengan keinginan kita.

Referensi: https://yudiwbs.wordpress.com/2018/03/29/ner-named-entity-recognition-dengan-anago-python-keras/

**PRETRAINED MODEL**

Model yang sudah dilakukan pelatihan oleh penulis adalah hasil pekerjaan dengan tim. Model tersebut sudah memuat deteksi pada 15 entitas yaitu:

1. PANJANG

label ini diberikan kepada seluruh informasi yang berhubungan dengan panjang, jarak dan tinggi. sebagai contoh:

1 meter, 1 cm, 2 mm, dua meter, 3 sentimeter, 3 senti meter, dll.

1. LUAS

label ini diberikan kepada seluruh informasi yang berhubungan dengan luas. sebagai contoh: 1 km2, 1 hektar, 2 hekto are, dst.

1. MASSA

label ini diberikan kepada seluruh informasi yang berhubungan dengan massa dan berat. sebagai contoh: 1kg, 1 gram, 2 kilogram, 2 kilo gram, dst.

1. SUHU

label ini diberikan kepada seluruh informasi yang berhubungan dengan suhu dan temperatur. sebagai contoh: 100 derajat celcius, 100 derajat fahrenheit, 1 °C, dst

1. WAKTU

label ini diberikan kepada seluruh informasi yang berhubungan dengan waktu. sebagai contoh: 1 tahun, 1 detik, 2jam, 2 jam, 2 dekade, dst.

1. KECEPATAN

label ini diberikan kepada seluruh informasi yang berhubungan dengan kecepatan. sebagai contoh: 1 m/s, 20 km per jam, 20 meter per detik, dst.

1. VOLUME

label ini diberikan kepada seluruh informasi yang berhubungan dengan volume sesuatu. sebagai contoh: 1 meter kubik, 2 dm3, 3 m3, dst.

1. UANG

label ini diberikan kepada seluruh informasi yang berhubungan dengan mata uang. rp 2 juta, USD 500, 2 juta rupiah, dst.

1. PER

label ini diberikan kepada seluruh informasi yang berhubungan dengan nama orang. sebagai contoh: joko widodo, bapak joko widodo, H. Jokowi, Kombespol Jokowi, dst.

1. LOC

label ini diberikan kepada seluruh informasi yang berhubungan dengan nama suatu tempat baik itu fiktif atau nyata selain nama negara. sebagai contoh: kota surabaya, desa adat merbabu, kecamatan tandes, kel keputih, taman safari prigen, air terjun kakek bodo, Benua afrika dst, amerika.

1. ORG

label ini diberikan kepada seluruh informasi yang berhubungan dengan organisasi atau institusi baik iku pemerintahan ataupun swasta. contoh: kementerian keuangan, pt. cahaya buana, pt angkasa pura, satpol pp, dst.

1. TIME

label ini diberikan kepada seluruh informasi yang berhubungan dengan hari tanggal dan waktu yg sudah jelas. contoh: januari 2020, senin 2 agustus 1945, 13.00 WIB, senin pukul 13.00 WIB, dst.

1. QTY

label ini diberikan kepada seluruh informasi yang berhubungan dengan kuantitas, angka atau segala sesuatu yang tidak dapat didefinisikan ke dalam poin 1-8. contoh: 2 mobil, 2 orang, 2 hewan, 2 buah permen, 2 jembatan, 200 pengungsi, 4 skala rikter, dst.

1. NEGARA

label ini diberikan kepada seluruh informasi yang berhubungan dengan nama negara. contoh: Indonesia, republik indonesia, USA, UAE, Amerika serikat, Malaysia, brunei darussalam, RRC, Republik Rakyat China dst.

1. EMAIL

label ini diberikan kepada seluruh informasi yang berhubungan dengan email. contoh: abc@gmail.com, blabla@digital.com, dst