



DATA MINING

DALAM NATURAL LANGUAGE PROCESSING





[HTTPS://WWW.LINKEDIN.COM/
IN/ABDAZZAMAJHARI/](https://www.linkedin.com/in/abdaazzamajhari/)



PROCTOR



ABDUL "AZZAM" AJHARI

PENGALAMAN KERJA

(2019-SEKARANG)

- DATA SCIENTIST DI BADAN SIBER DAN SANDI NEGARA (BSSN)

(2018-2019)

- IT SYSTEM DEVELOPMENT OFFICER DI PT SINAR NIAGA SEJAHTERA
(DISTRIBUTOR PT GARUDAFOOD TBK).

PENDIDIKAN

- MAGISTER TEKNIK INFORMATIKA, BINUS GRADUATE PROGRAM.
PENJURUSAN DATA SCIENCE. (2020-2022)
- SARJANA TEKNIK INFORMATIKA, UNIVERSITAS BUDI LUHUR.
PENJURUSAN KRIPTOGRAFI DAN KEAMANAN SIBER. (2014-2018)

PENGHARGAAN

JUARA #2 TIM MATASIBER PADA INDIGO HACKATHON FESTIVAL 2021
KATEGORI PENINGKATAN TRANSPORTASI DAN LOGISTIK TENTANG
PENDETEKSIAN ANOMALI PADA PESAWAT SEKTOR PENERBANGAN.

DATA MINING



“Data mining adalah proses mengidentifikasi pola atau struktur dan pengetahuan dari data yang dikumpulkan.”

(J. Roiger, 2016)

BUSINESS UNDERSTANDING

Tentukan tujuan bisnisnya
(masalah dan solusi).



DATA ACQUISITION AND UNDERSTANDING

	created_at	tweet	user
0	2022-03-19 14:05:45+00:00	b'@HASapardan Lanjut ! Setuju ajukan hak angke...	Mbah Tokok
1	2022-03-19 13:35:44+00:00	b'Kemarin waktu minyak goreng langka. saya dar...	Tacipi City
2	2022-03-19 13:22:45+00:00	b'Puan Murka Harga Minyak Goreng Belum Normal:...	Ayun hamunta
3	2022-03-19 13:20:43+00:00	b'Harga minyak goreng curah pada tingkat agen ...	Media Indonesia
4	2022-03-19 13:20:05+00:00	b'Menolak lupa!!!\nDulu BBM naik, demo berjili...	fian baru
...
8457	2022-04-16 04:06:36+00:00	b'Cek Ketersediaan Minyak Goreng, Kapolres Pas...	Ini Balikpapan
8458	2022-04-16 04:06:31+00:00	b'79,3 Persen Masyarakat Puas Atas Kinerja Pre...	dwi kumalasari
8459	2022-04-16 04:06:21+00:00	b'Babinsa Koramil 10/Prembung melaksanakan pend...	Kodim 0709/Kebumen
8460	2022-04-16 04:05:51+00:00	b'Pemerintah Pastikan Stok Minyak Goreng Aman ...	dwi kumalasari
8461	2022-04-16 04:05:48+00:00	b'@CaryaOkada Minyak goreng 50 rb nya kakack'	Dedek Gupi

8462 rows × 3 columns



Pelajari tipe dan kualitas datanya. .

DATA PREPARATION/PREPROCESSING

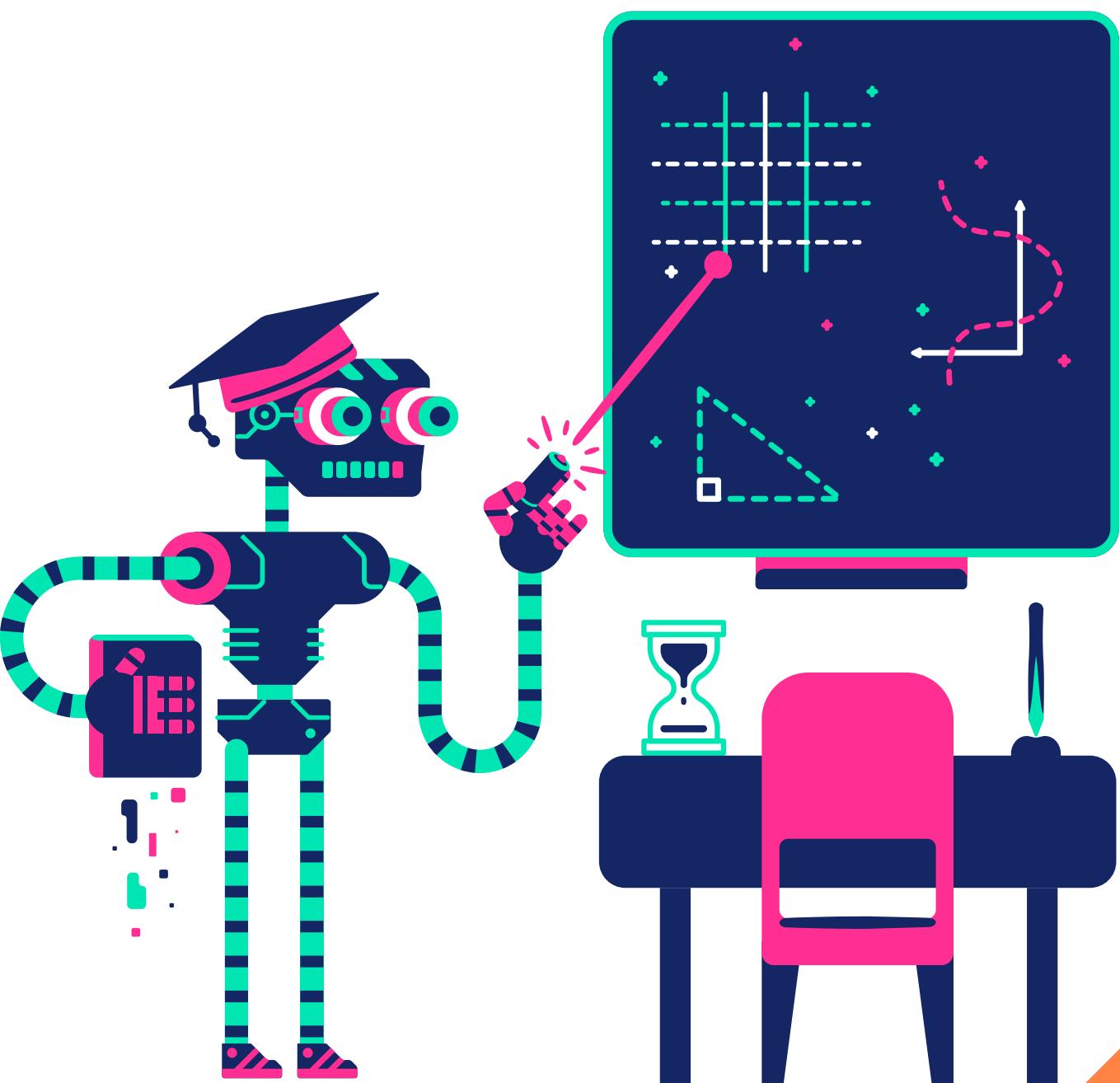


Tahapan pembersihan data.

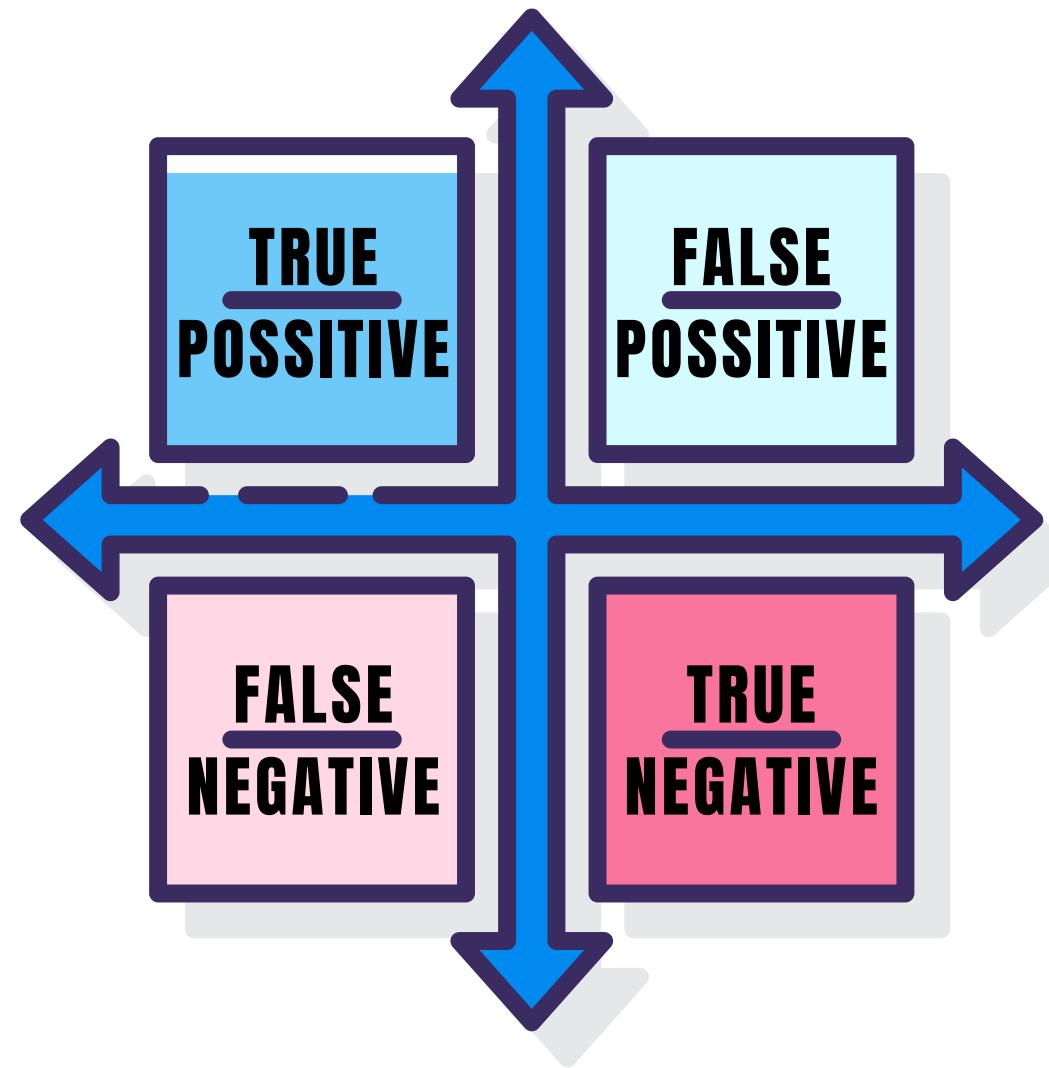
- Tokenizing
- RegEx
- etc.

MODELING

Tentukan model Algoritma yang ingin digunakan untuk melatih data.

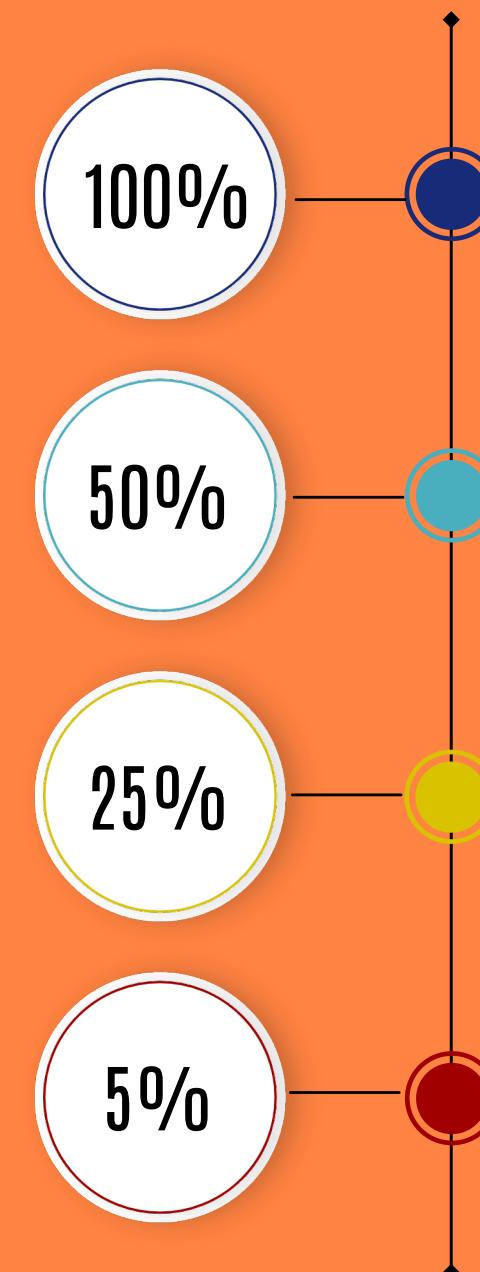


EVALUATION



Berikan hasil performa pengujian dan tahapan implementasinya.
Dapat menggunakan Confusion Matrix.

REPORTING



Mendokumentasikan
seluruh kegiatan yang
teknikal. Nilai tambah
apabila dibuat ke dalam
bahasa **manajerial**.

study
case



pythonTM