

Steps:

Standard Kompetensi Kerja Nasional: KepMen Ketenagakerjaan No 299 thn 2020

FUNGSI KUNCI	FUNGSI UTAMA	FUNGSI DASAR
Menganalisis Kebutuhan (Requirements) Organisasi	<i>Business Understanding</i>	1. Menentukan objektif bisnis 2. Menentukan tujuan teknis 3. Membuat rencana proyek
	<i>Data Understanding</i>	4. Mengumpulkan data 5. Menelaah data 6. Memvalidasi data
Mengembangkan model	<i>Data Preparation</i>	7. Memilah data 8. Membersihkan data 9. Mengkonstruksi data 10. Menentukan Label Data 11. Mengintegrasikan data
	<i>Modeling</i>	12. Membangun skenario pengujian 13. Membangun model
	<i>Model Evaluation</i>	14. Mengevaluasi hasil pemodelan 15. Melakukan review proses pemodelan
Menggunakan model yang dihasilkan	<i>Deployment</i>	16. Membuat rencana deployment model 17. Melakukan deployment model 18. Melakukan rencana pemeliharaan 19. Melakukan pemeliharaan
	<i>Evaluation</i>	20. Melakukan review proyek 21. Membuat laporan akhir proyek

Case: BPPJS Hackaton

- *Fraud Detection (Binary Classification)*
- *Case and Cost Prediction (Regression Problem)*

Case: BPPJS Hackaton (1)

- ***Fraud Detection (Binary Classification)***

- Mengembangkan sebuah model data mining untuk melakukan prediksi potensi terjadinya fraud pada klaim pelayanan Rumah Sakit berdasarkan dataset train yang terdiri dari 200217 observasi dan 53 variable yaitu:
 - *visit_id* = id kunjungan
 - *kdkc* = kode wilayah kantor cabang BPJS Kesehatan
 - *dati2* = kode kabupaten/kota
 - *typeppk* = kode tipe Rumah Sakit
 - *jkpst* = jenis kelamin peserta JKN-KIS
 - *umur* = umur peserta saat mendapatkan pelayanan rumah sakit
 - *jnspelsep* = tingkat pelayanan; 1:rawat inap; 2. rawat jalan
 - *los* = lama peserta dirawat di rumah sakit
 - *cmg* = klasifikasi CMG (Case Mix Group)
 - *severitylevel* = tingkat urgensi
 - *diagprimer* = diagnosa primer
 - *dx2_..._...* = diagnosa sekunder
 - *proc.._...* = kode kelompok procedure
 - *label* = flag fraud; 1:fraud; 0:tidak fraud

Case: BPPJS Hackaton (2)

- ***Case and Cost Prediction (Regression Problem)***

- *Mengembangkan sebuah model data mining untuk melakukan prediksi jumlah kasus dan unit cost pada sebuah daerah akibat penambahan Rumah Sakit kerja sama berdasarkan dataset train yang terdiri atas 57971 observasi dan 36 variable yaitu:*

- *row_id : id*
- *tglpelayanan : periode bulan pelayanan di rumah sakit*
- *kddati2 = Kode kabupaten/kota*
- *tkp = tingkat pelayanan; 30:rawat jalan; 40:rawat inap*
- *peserta = jumlah peserta aktif pada kabupaten/kota periode tersebut*
- *a,b,c, ... ,sd = tipe rumah sakit yang melayani peserta JKN-KIS*
- *case : jumlah kunjungan rumah sakit*
- *unit_cost = jumlah biaya pelayanan rumah sakit*

Case: BPPJS Hackaton (Requirement)

- Kedua kategori tersebut harus dibagi menjadi *data training* dan *data validation* berbeda
- Classification problem:
 - *Precision* > 0.60
 - *Accuracy* > 0.60
 - *Recall* > 0.65
- Regression problem:
 - Kasus kunjungan: *MAE* < 900 , *MAPE* $< 90\%$
 - Biaya: *MAE* < 97000 , *MAPE* $< 70\%$

Thank You

Colin Powell

"A dream does not become reality through magic; it takes sweat, determination, and hard work."