



INSTITUT TEKNOLOGI SUMATERA

Terrarium

BPJS HEALTHKATON **2022**

Machine Learning

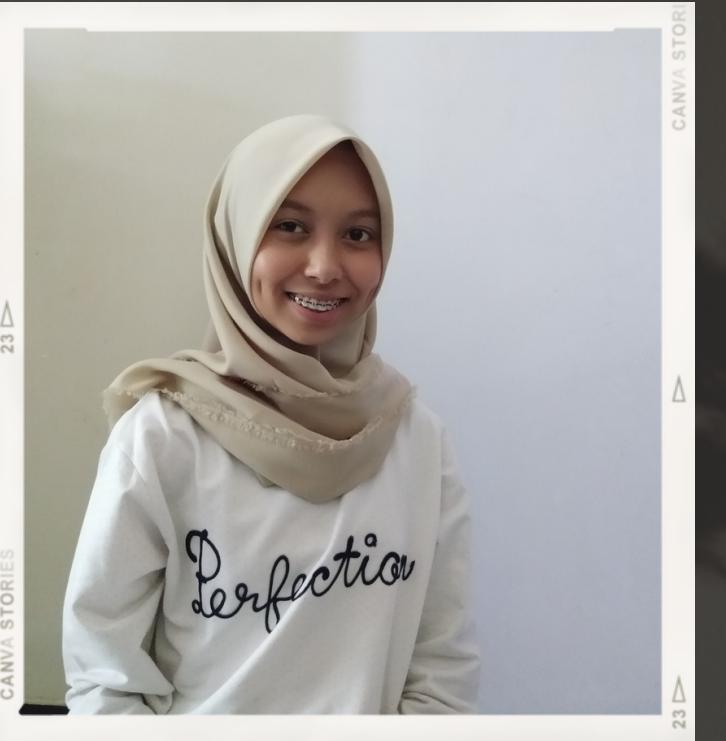
Anggota Kami



Lanang Almasyuri
Anggota Tim



M. Dhoni Apriyadi
Ketua Tim



Shahnaz Salsabila I
Anggota Tim

Terrarium

Overview

- ❖ Introduction
- ❖ Features Selection
- ❖ Pemilihan Algoritma
- ❖ Data Preparation
- ❖ Features Engineering
- ❖ Hasil Pengembangan
- ❖ Kemudahan Implementasi

Terrarium

Data Preparation

Page 04

Hapus baris data yang tidak memiliki format yang sesuai

Mengisi Data yang Nilainya Kosong

Data Procedure

Data Utama

Data Diagnosa

Merubah Tipe Data

Data Utama

PENGISIAN PADA DATA YANG HILANG

Data Nominal



'N'

Data Ratio



median

Data Preparation

MENGUBAH TIPE DATA

Int64

Int8

Int16

Int32

tanggal (string/object)

datetime

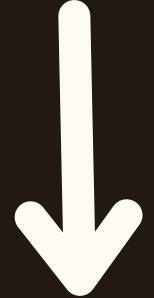
Data Procedure

Aturan penulisan kode procedure berdasarkan ICD-9-CM

4 digit angka

2 digit pertama : kode dasar.
karakter setelah desimal menunjukkan
subkategori procedure lebih spesifik

titik desimal antara
digit ke 2 dan 3



Menghapus baris yang formatnya tidak sesuai
dengan aturan penulisan kode procedure

Abjad

simbol seperti
' _ '

Data Procedure

Mengambil 2 digit pertama kode procedure

90.59 → 90

69.01 → 69

88.78 → 88

Menggabungkan procedure berdasarkan id

id	procedure
6	90
6	69
6	88
27	90
27	88

id	procedure
6	90, 69, 88
27	90, 88

Data Diagnosa

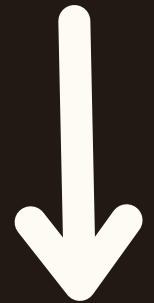
Aturan penulisan kode diagnosa berdasarkan ICD-10

3 - 7
karakter

titik desimal antara
karakter ke 3 dan 4

3 karakter pertama : kode dasar.
karakter setelah titik desimal menunjukkan
subkategori diagnosa lebih spesifik

diawali 1 huruf
abjad dengan
huruf besar



Filter kode diagnosa yang mengandung
abjad

Mengubah abjad di semua kode diagnosa
menjadi huruf besar

Data Diagnosa

Mengambil 3 karakter pertama kode diagnosa

006.9 → 006

M10.99 → M10

T14.70 → T14

Data Preparation

Menggabungkan diagnosa dan levelid berdasarkan id

id	diagnosa	levelid
11426954	M79	2
11426954	R51	1
11426954	JO6	2
11426957	M79	1
11426957	K30	2
11426957	R10	2

id	diagnosa	levelid
11426954	M79, R51, JO6	1,2
11426957	M79, K30, R10	1,2



Features Selection

Terrarium

Data Utama

22 features

id	id_peserta	dati2	typefaskes	usia	jenkel	pisat	tgldatang	tglpulang	label
----	------------	-------	------------	------	--------	-------	-----------	-----------	-------	-------



7 features

id	dati2	tgldatang	tglpulang	politujuan	biaya	label
----	-------	-----------	-----------	------------	-------	-------

Features Selection

Data Utama

Memilih politujuan yang sering muncul pada data yang diberi label 1 atau memiliki potensi inefisiensi pada klaim

Features Selection

Memilih politujuan yang muncul lebih dari 1500 kali

politujuan	sum
N	53563
IGD	20800
MAT	19288
INT	18755
BED	15040

13 politujuan

*preview

Data Procedure dan Diagnosa

Memilih kode procedure dan kode diagnosa yang sering muncul pada data yang diberi label 1 atau memiliki potensi inefisiensi pada klaim

Memilih kode procedure yang muncul lebih dari 1000 kali

15 procedure	procedure	sum
	99	26634
	89	19288
	90	18755
	93	15040

*preview

Memilih kode diagnosa yang muncul lebih dari 2000 kali

21 diagnosa	diagnosa	sum
	PO3	13219
	H52	8112
	E11	5671
	K30	4716

*preview

Features Selection



Features Engineering

Terrarium

Data Utama



Menambah fitur hari untuk menentukan lamanya (dalam hari) pasien dirawat

Menghitung selisih
tanggal datang dan
tanggal pulang



Menghapus kata 'days'
pada isi feature



Mengubah tipe data
menjadi int16

Menghapus feature tgldatang dan tglpulang

Data Utama

Memberikan variabel dummy pada setiap setiap data berdasarkan
politujuan yang telah didapatkan pada features selection

id	label	IGD	MAT	INT	HDL
165666	0	0	0	0	0
166042	0	0	0	1	0
168937	0	0	1	0	0

Data Procedure

Memberikan variabel dummy pada setiap setiap data berdasarkan kode procedure yang telah didapatkan pada features selection

id	99	89	90	56
1	0	0	1	0
3	0	0	1	0
4	0	0	1	0

Data Diagnosa

Memberikan variabel dummy pada setiap setiap data berdasarkan kode diagnosa yang telah didapatkan pada features selection dan levelid

id	H52	E11	K30	J06	1	2
2	0	0	0	0	1	0
3	1	0	0	0	1	0
4	0	0	0	0	1	0
5	0	0	1	0	1	0

MERGING

Data Utama + Procedure + Diagnosa

id	dati2	biaya	label	hari	N	IGD	MAT	INT	R1O	LO2	1	2
1010828	17	10628400	0	3	1	0	0	0	0	0	0	1	1
168937	17	187300	0	0	0	0	1	0	0	0	0	1	1

56 Columns



Pemilihan Algoritma

Terrarium

Pemilihan Algoritma





Hasil Perkembangan

Terrarium

Hasil Perkembangan

Data Test

Sensitivity : 0.5094

Precision : 0.7226

F1-Score : 0.5976

Specificity : 0.9973

Data Validasi 3

Sensitivity : 0.495

Precision : 0.685

F1-Score : 0.5747

Specificity : 0.997



Kemudahan Implementasi

Terrarium

Kemudahan Implementasi

01

Save Model Fit With Pickle

02

Load and Deploy With Flask

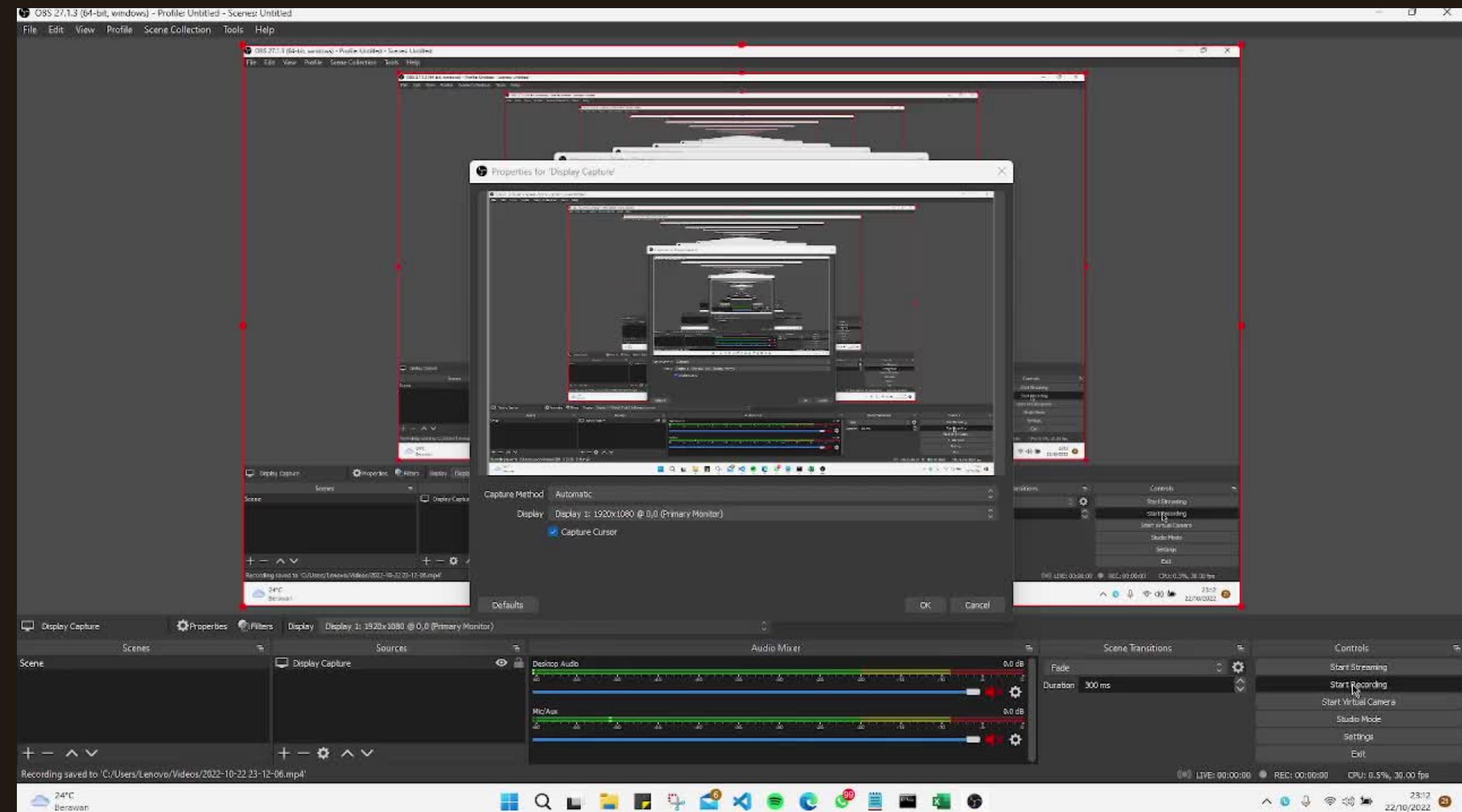
The screenshot shows a web application titled "Fraud Detection [Terrarium]". The interface is divided into several sections:

- Main Data:** Contains a "Choose File" button with the message "No file chosen".
- Procedure Data:** Contains a "Choose File" button with the message "No file chosen".
- Diagnosis Data:** Contains a "Choose File" button with the message "No file chosen" and a "Submit Data" button.
- Hasil Prediksi:** This section is currently empty.

Below the form is a file browser table:

Name	Date modified	Type	Size
__pycache__	22/10/2022 22:33	File folder	
datasample100k	22/10/2022 23:16	File folder	
templates	22/10/2022 22:32	File folder	
thedata	22/10/2022 23:12	File folder	
app.py	22/10/2022 22:39	PY File	3 KB
bpjs_final_model.sav	22/10/2022 14:54	SAV File	36,956 KB
diag_prev.csv	22/10/2022 23:10	Microsoft Excel C...	1 KB
main_prev.csv	22/10/2022 23:10	Microsoft Excel C...	1 KB
prediction.csv	22/10/2022 23:12	Microsoft Excel C...	1 KB
preprocessing.py	22/10/2022 21:14	PY File	8 KB
proc_prev.csv	22/10/2022 23:10	Microsoft Excel C...	1 KB

Kemudahan Implementasi



Thank
You!

TERRARIUM || MACHINE LEARNING