

LAPORAN PROYEK DATA MINING
Classification of Tuberculosis Disease from 2015-2021 with
Artificial Neural Network (ANN)



Disusun Oleh:

12S20006	Nadya Dioranta Tambunan
12S20022	Agnes Veronika Sihombing
12S20044	Hanna Dhea Christi Sihombing

PROGRAM STUDI SARJANA SISTEM INFORMASI
FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL

2023

DAFTAR ISI

DAFTAR GAMBAR.....	3
DAFTAR TABEL	4
BAB I BUSINESS UNDERSTANDING.....	5
1.1 Menentukan Objektif Bisnis.....	5
1.2 Menentukan Tujuan Bisnis.....	5
1.3 Membuat Rencana Proyek.....	5
BAB II DATA UNDERSTANDING	8
2.1 Mengumpulkan Data	8
2.2 Menelaah Data.....	8

DAFTAR GAMBAR

DAFTAR TABEL

Tabel 1 Jadwal Pelaksanaan Proyek.....	6
Tabel 2 Spesifikasi development environment.....	7

BAB I

BUSINESS UNDERSTANDING

Bab ini memperjelas pemahaman konten dan kebutuhan operasi data mining dari sudut pandang bisnis. Langkah-langkahnya mencakup penetapan tujuan bisnis, pemahaman mendalam terhadap situasi bisnis, serta interpretasi tujuan bisnis atau sasaran bisnis menjadi tujuan spesifik dalam konteks operasi data mining.

1.1 Menentukan Objektif Bisnis

Badan Penyelenggara Jaminan Sosial atau BPJS Kesehatan adalah badan hukum publik yang memiliki tanggung jawab terhadap penyediaan jaminan serta layanan kesehatan bagi masyarakat atau pesertanya. BPJS Kesehatan telah berhasil mengelola lebih dari 300 miliar baris data yang tersimpan dalam Big Data BPJS Kesehatan. Data ini mencakup informasi kepesertaan dan pelayanan kesehatan yang diperoleh dari masyarakat yang menggunakan Jaminan Kesehatan Nasional (JKN) untuk mengakses fasilitas pelayanan kesehatan. Sebagai upaya untuk mendorong partisipasi dan inisiatif masyarakat dalam menganalisis data tersebut, BPJS Kesehatan kemudian meluncurkan data sampel BPJS Kesehatan. Pada data sampel yang terbaru, terdapat penambahan data sampel kontekstual penyakit Tuberkulosis yang masih merupakan masalah yang signifikan di Indonesia pada saat ini. Data sampel ini merupakan pengembangan dari data sampel kontekstual pada tahun 2021 yang mengeluarkan data sampel penyakit Diabetes Mellitus.

Pada proyek ini, akan melakukan analisis dan klasifikasi pada data fasilitas pelayanan kesehatan peserta BPJS yang mengalami penyakit Tuberkulosis (TB). Objektif utama dalam pengerjaan proyek ini adalah melakukan prediksi untuk pasien penyakit Tuberkulosis (TB).

1.2 Menentukan Tujuan Bisnis

Tujuan dari proyek ini yaitu ingin membangun sebuah model yang efisien terkait pengelompokkan pelayanan kesehatan BPJS untuk menganalisis dan memahami fasilitas pelayanan kesehatan pada penyakit Tuberkulosis.

1.3 Membuat Rencana Proyek

Table dibawah memperlihatkan tahapan perencanaan yang dilakukan dalam mencapai tujuan *data mining* dan tujuan bisnis dalam “*Classification of Tuberculosis Disease from 2015-2021*”.

Tabel 1 Jadwal Pelaksanaan Proyek

Tahapan	Kegiatan	Waktu
Pemilihan Case/Kasus	Menentukan kasus dan algoritma.	5 hari
<i>Business Understanding</i>	Menentukan objektif bisnis, tujuan bisnis, dan membuat project plan.	3 hari
<i>Data Understanding</i>	Data collection, mendeskripsikan data, melakukan eksplorasi data dan memvalidasi data.	4 hari
<i>Data Preparation</i>	Memilih dat, membersihkan data, membangun data, menggabungkan data dan mengintegrasikan data.	5 hari
<i>Modelling</i>	Memilih teknik pemodelan, membuat Test Design, membangun model, dan menilai model	5 hari
<i>Evaluation</i>	Mengevaluasi hasil dan melakukan review proses pemodelan	4 hari
<i>Deployment</i>	Membuat plan deployment, Monitoring and Maintenance Plan danmembuat laporan akhir proyek.	7 hari

Untuk mendukung pengerjaan setiap tahapan proses pada proyek ini, berikut adalah spesifikasi development environment pada proyek ini:

Tabel 2 Spesifikasi development environment

Tools	<ul style="list-style-type: none">- Jupyter Notebook- Google Collab- Visual Studio Code
Bahasa Pemrograman	Python3
Algoritma	Artificial Neural Network (ANN)
Web Interface	<i>Flask Python</i>
Deployment Cloud	-

BAB II

DATA UNDERSTANDING

Pada bab ini akan dijelaskan mengenai data collection, mendeskripsikan data, melakukan eksplorasi data dan memvalidasi data.

2.1 Mengumpulkan Data

Langkah pertama yang dilakukan pada *data understanding* adalah pengumpulan data dalam persiapan awal dalam melakukan proyek ini yang akan digunakan pada kasus *Classification of Tuberculosis Disease from 2015-2021 with Artificial Neural Network (ANN)*. Dataset yang digunakan adalah data BPJS Kesehatan yaitu data TB20152021_fkrtl.dta, TB20152021_fkrtldxsekunder.dta, TB20152021_fktponkapitasi.dta yang bersifat statis dan memiliki format file dta yang berasal dari website BPJS Kesehatan.

2.2 Menelaah Data

Pada tahap ini merupakan tahap untuk mempelajari, mengkaji data yang dimiliki. Dataset yang dimiliki akan digunakan untuk melakukan prediksi terjadinya fraud atau tidak fraud pada pelayanan rumah sakit.