



PROJECT
DATA MINING

CLASSIFICATION OF TUBERCULOSIS DISEASE FROM 2019-2021 USING ARTIFICIAL NEURAL NETWORK (ANN)

Group : 04

a Mining Project.

BPJS Classification System

Analisis dan klasifikasi pada data fasilitas kesehatan peserta BPJS yang mengalami Tuberkulosis (TB).

Check Now

OUR TEAM

1



**Nadya Dioranta
Tambunan**

12S20006

2



**Agnes Veronika
Sihombing**

12S20022

3



**Hanna Dhea C
Sihombing**

12S20044

Table of Content

- 01**
Business Understanding
- 02**
Data Understanding
- 03**
Data Preparation
- 04**
Modelling
- 05**
Evaluation
- 06**
Deployment

ing Project.

BPJS

Nomor Peserta *

Kabupaten/Kota Tempat Tinggal Peserta *

Kabupaten/Kota Fasilitas Kesehatan Peserta Terdaftar *

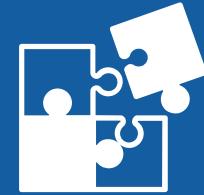
Tipe faskes *

Nama Diagnosis *

All fields marked with an asterisk (*) are required

Predict

Business Understanding



OBJEKTIF BISNIS

Kurangnya informasi dan ketidaksesuaian layanan FKTP BPJS menyebabkan ketidakpuasan peserta Tuberkolosis. Perlu dilakukan klasifikasi fasilitas layanan peserta BPJS dengan Tuberkolosis untuk periode 2019-2021.



TUJUAN BISNIS

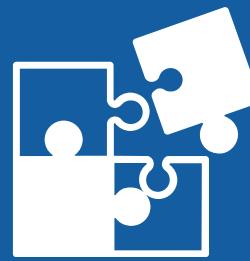
Tujuan dari proyek ini yaitu ingin membangun sebuah model yang efisien terkait pengelompokan pelayanan kesehatan BPJS untuk menganalisis dan memahami fasilitas pelayanan kesehatan pada penyakit Tuberkulosis.



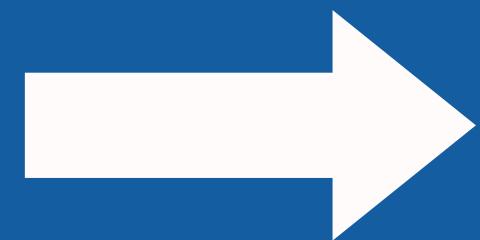
RENCANA PROSES

Proyek dilakukan selama 7 minggu, dimulai dari 20 November 2023 - 5 Januari 2024.

Data Understanding

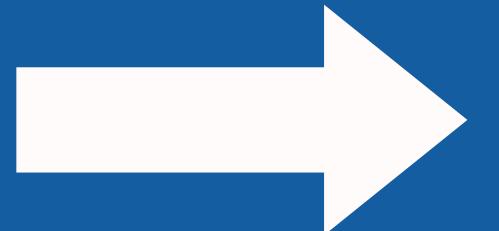


MENGUMPULKAN DATA



- Sumber dataset : BPJS 2019–2022
- Total file data : 6
- Format dataset : dta

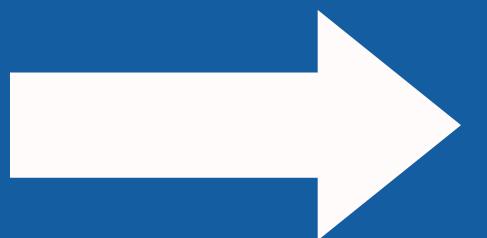
Data Understanding



MENELAAH DATA

1. **TB20152021_fktpnonkapitasi.dta**
Atribut : 21, Baris : 5, Kolom : 21
2. **TB2019_kepesertaan.dta**
Atribut : 18, Baris : 94.966, Kolom : 18
3. **TB2020_kepesertaan.dta**
Atribut : 18, Baris : 94.966, Kolom : 18
4. **TB2021_kepesertaan.dta**
Atribut : 18, Baris : 94.966, Kolom : 18

Data Understanding

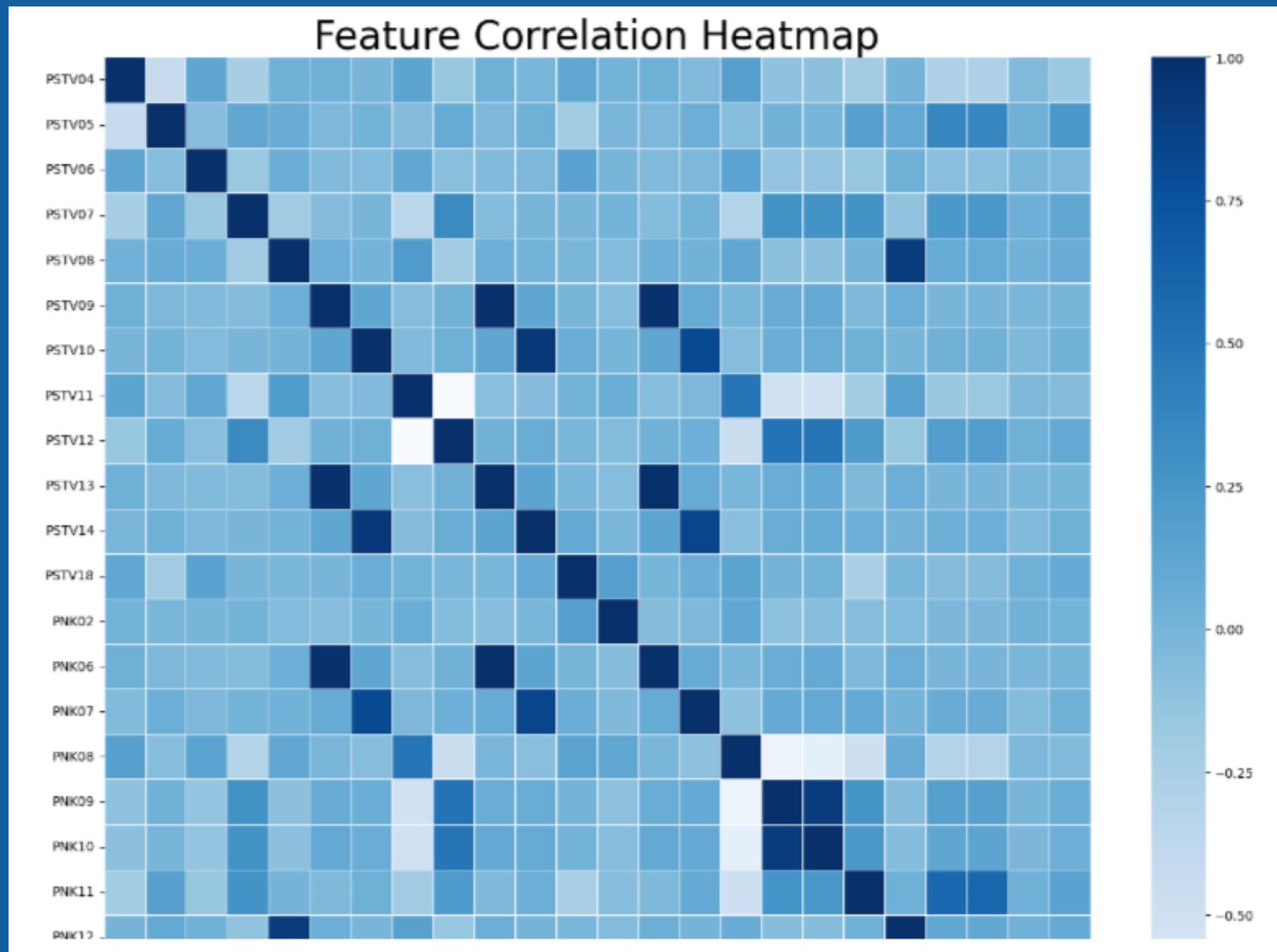


MENELAAH DATA

Rangkuman Data Understanding :

- Memeriksa 5 data pertama
- Memeriksa Shape Data
- Memeriksa Columns Data
- Memeriksa Data Info
- Memeriksa Data Atribut Types
- Memeriksa Deskripsi Data
- Memeriksa Missing Values
- Memeriksa Data Unique
- Memeriksa Correlation

Data Correlation



Setelah melakukan korelasi pada keempat data yang digunakan, maka data akan di merge.

Data Understanding



MEMVALIDASI DATA



- Missing Value
Jumlah : 102036 pada atribut PSTV18
- Duplicated
Jumlah : 1461

DATA PREPARATION

01 Memilih Data dan Membersihkan Data

- Menghapus missing value & duplicated

Setelah melakukan merge data pada 4 data yang digunakan, maka objek data yang tidak sesuai akan dihapus. Objek data yang memiliki missing value adalah atribut 'PSTV18'. Data yang duplicated juga akan dihapus.

```
# handling missing value
df['PSTV18'] = df['PSTV18'].fillna('Belum Meninggal')

[ ] df['PSTV18']

0      Belum Meninggal
1      Belum Meninggal
2      Belum Meninggal
3      Belum Meninggal
4      Belum Meninggal
...
109474    Belum Meninggal
109475    Belum Meninggal
109476    Belum Meninggal
109477    Belum Meninggal
109478    Belum Meninggal
Name: PSTV18, Length: 109479, dtype: object
```

```
✓ Check Duplicated

[ ] df = df.duplicated()
df.sum()

1461

✗ Drop Duplicated

[ ] df = df.drop_duplicates(keep='first', inplace=False)

[ ] df = df.duplicated()
df.sum()

0
```

DATA PREPARATION

O1 Memilih Data dan Membersihkan Data

- Mengeliminasi Data yang Tidak Relevan

Atribut yang tidak relevan dan dieliminasi adalah ['PSTV18', 'PSTV17', 'PSTV16', 'PSTV11', 'PSTV12', 'PSTV08', 'PSTV13', 'PSTV07', 'PSTV04', 'PSTV06', 'PNK13A', 'PNK12', 'PSTV05', 'PNKO6', 'PNKO9', 'PNKO8', 'PNK10'].

DATA PREPARATION

O2 Merekonstruksi Data

- Transformasi

Mengubah kategori object menjadi numerikal.

Atribut yang digunakan adalah PNK11 dan isinya diubah menjadi numerikal yaitu:

'RITP' = 2

'RJTP' = 1

'PROMOTIF' = 0

DATA PREPARATION



03 Menentukan Label Data

Dataset yang digunakan sudah memiliki label data yang dapat langsung digunakan untuk pemodelan nantinya, sehingga tahapan pelabelan data tidak perlu dilakukan lagi.

DATA PREPARATION

04 Mengintegrasikan Data

Terdapat 4 data dengan format dta yang diintegrasikan dan menghasilkan satu dataset. Namun pertama sekali, data yang di merged adalah data kepesertaan dari tahun 2019–2021.

Kemudian hasil merge data kepesertaan digabungkan lagi dengan data fktpnonkapasiti, namun dikarenakan pada data kepesertaan sudah terdapat 3 atribut ('PSTV01', 'PSTV02', 'PSTV15'), maka pada data fktpnonkapasiti tidak perlu dimasukkan 3 kolom atribut tersebut.

MODELING

01 Skenario Pengujian

Model klasifikasi yang diharapkan berdasarkan yang dibangun memenuhi syarat berupa nilai dari:

Precision > 0.6

Accuracy > 0.6

Recall > 0.6

Pembagian Dataset yaitu:

Data Train : 70%

Data Test : 30%

MODELING

O2 Algoritma

Artificial Neural Network adalah salah satu metode Deep Learning yang bekerja seperti jaringan saraf manusia. Kategori ANN didasarkan pada metode supervised dan unsupervised.

Bentuk sederhana dari ANN adalah persepsi yang terdiri dari satu dengan dua masukan dan satu keluaran. Persepsi digunakan untuk klasifikasi data yang diubah menjadi dua kelas terpisah.

MODELING

O3 Build Model

Dalam pembuatan/pembangunan model ini, kami melakukan tahapan sebagai berikut:

Import library, import dataset, defenisi x dan y, membagi data train dan data test, build model, copile model, dan train.

Dengan epoch 10 dan batch size 5

Dan Compile model dengan adam, binary cross dan metrics akurasi.

EVALUATION

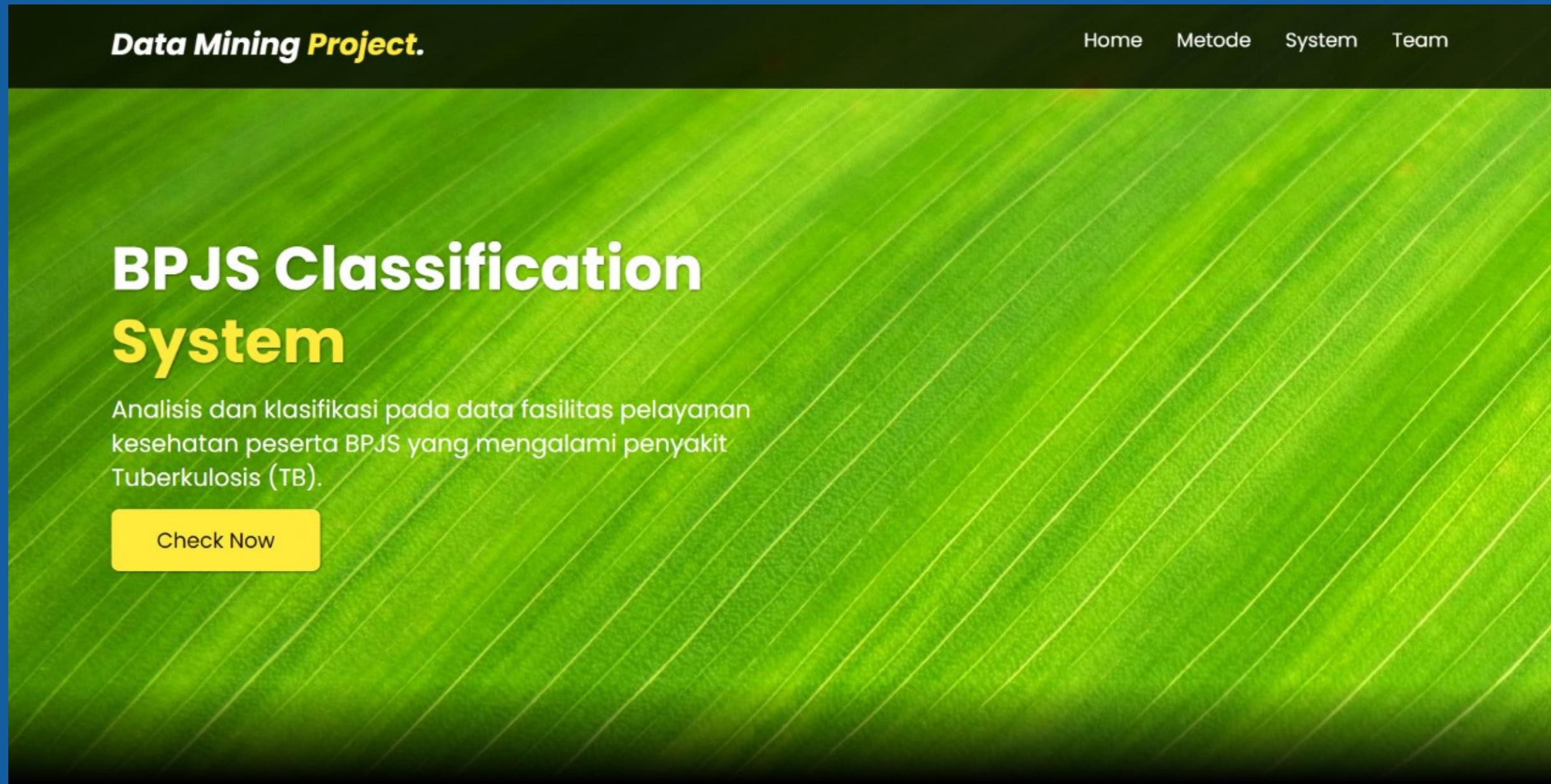
Hasil evaluasi model yang didapatkan, yaitu:

- Test Accuracy : 0.70 / 70%
- Test Precision : 76%
- Test Recall : 98%
- F1 Score : 85%

DEPLOYMENT

Deployment digunakan menggunakan Render.com yang dihubungkan menggunakan Github Repository.

- Halaman Utama



DEPLOYMENT

- Halaman Metode

The screenshot shows a website with a dark blue header and a light green main content area. The header contains the text "Data Mining Project." in white, followed by a navigation menu with links "Home", "Metode", "System", and "Team". The main content area has a title "Metode Kami" in bold yellow text. Below it, there is a section titled "Artificial Neural Network" in bold white text. A descriptive paragraph follows, explaining that Artificial Neural Network is a type of Deep Learning that works like a human neural network, using supervised and unsupervised methods. It then lists three points under the heading "Arsitektur dari ANN adalah sebagai berikut:".

Data Mining Project.

Home Metode System Team

Metode Kami

Artificial Neural Network

Artificial Neural Network adalah salah satu metode Deep Learning yang bekerja seperti jaringan saraf manusia. Kategori ANN didasarkan pada metode supervised dan unsupervised. Bentuk sederhana dari ANN adalah persepsi yang terdiri dari satu dengan dua masukan dan satu keluaran. Persepsi digunakan untuk klasifikasi data yang diubah menjadi dua kelas terpisah.

Arsitektur dari ANN adalah sebagai berikut:

- a. Lapisan Input (Input Layer): Menerima nilai input
- b. Lapisan tersembunyi (Hidden Layer): Satu set neuron antara lapisan input dan output.
- c. Lapisan Output (Output layer): Memiliki satu neuron dan outputnya berkisar antara 0 dan 1 atau lebih besar dari 0 dan lebih kecil dari 1.

DEPLOYMENT

- Halaman Sistem

Data Mining Project.

Home Metode System Team

BPJS Classification using ANN

Fill-up the form to predict

Nomor Peserta *

Kabupaten/Kota Tempat Tinggal Peserta *

Kabupaten/Kota Fasilitas Kesehatan Peserta Terdaftar *

Tipe faskes *

Nama Diagnosis *

All fields marked with an asterisk (*) are required

predict

Result:

DEPLOYMENT

- Halaman Team

Data Mining Project.

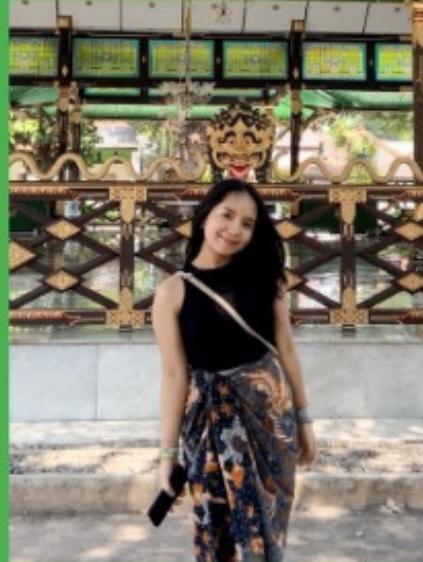
Home Metode System Team

Our Team

Kami dari kelompok 4 Project Data Mining yang melakukan project akhir mata kuliah untuk melakukan klasifikasi case pada data sampel BPJS 2015-2021



**Agnes Veronika
Sihombing**
12S20022



**Hanna Dhea Christi
Sihombing**
12S20044



**Nadya Dioranta
Tambunan**
12S20006



Thank you