

BAB 3

METODE PENELITIAN

Penelitian ini menggunakan metode penelitian klasifikasi dengan menggunakan *Naïve Bayes Classification* sebagai algoritma untuk mengklasifikasikan tingkat tekanan mental pada mahasiswa tingkat akhir studi di Fakultas Teknik Dan Teknologi Informasi. Penelitian ini terdapat beberapa tahapan metode penelitian diantaranya sebagai berikut:

3.1 BAHAN DAN ALAT PENELITIAN

Bahan yang dibutuhkan pada penelitian adalah item penilaian untuk kuesioner, *dataset* akan diperoleh melalui penyebaran kuesioner kepada mahasiswa tingkat akhir studi di Fakultas Teknik Dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta. Indikator penyusunan kuesioner pada penelitian ini akan didasarkan pada metode DASS – 42 untuk mengidentifikasi tingkat depresi, kecemasan dan stres, serta untuk sistem penilaian berdasarkan skala likert. Berikut ini merupakan item penilaian berbasis DASS – 42 yang dapat dilihat pada Tabel 3.1.

Tabel 3.1 Item Penilaian kuesioner DASS-42

Q	Item Penilaian
Q1	Saya sangat mudah merasa marah dikarenakan hal – hal kecil(sepele)
Q2	Saya sering kali bereaksi secara berlebihan dalam situasi tertentu
Q3	Saya sulit untuk bersantai atau merasa rileks
Q4	Saya cenderung merasa kesal
Q5	Saya cenderung cemas karena merasa kehabisan energi
Q6	Saya cenderung merasa tidak sabar
Q7	Saya rentan tersinggung
Q8	Saya memiliki kesulitan untuk beristirahat dengan baik
Q9	Saya rentan terhadap emosi marah

Q	Item Penilaian
Q10	Saya cenderung merasa kesulitan untuk kembali fokus setelah mengalami gangguan
Q11	Saya cenderung kesulitan untuk berkonsentrasi apabila ada gangguan
Q12	Saya cenderung merasa tegang
Q13	Saya tidak dapat menerima hal-hal yang menghambat penyelesaian aktivitas
Q14	Saya rentan terhadap perasaan gelisah
Q15	Saya seringkali merasa kering di mulut
Q16	Saya mengalami gangguan pernapasan seperti napas cepat atau kesulitan bernapas
Q17	Saya seringkali mengalami rasa lemah pada bagian tubuh tertentu
Q18	Saya seringkali merasa cemas yang berlebih pada kondisi tertentu, meskipun bisa merasa lega setelah kondisi tersebut berakhir
Q19	Saya seringkali merasa kelelahan secara terus menerus
Q20	Saya seringkali mengalami keringat berlebih tanpa alasan yang jelas
Q21	Saya cenderung merasa takut tanpa adanya alasan yang jelas
Q22	Saya cenderung kesulitan dalam menelan makanan atau minuman
Q23	Jantung saya sering berdenyut dengan cepat tanpa ada aktivitas fisik
Q24	Saya cenderung panik dalam kondisi tertentu
Q25	Saya cenderung takut terhambat oleh tugas-tugas yang tidak biasa
Q26	Saya cenderung merasa ketakutan tanpa adanya alasan yang jelas
Q27	Saya merasa khawatir terhadap situasi ketika saya mengalami panik atau malu
Q28	Saya seringkali mengalami gemetar
Q29	Saya kesulitan melihat sisi positif dari suatu situasi
Q30	Saya merasa tidak memiliki energi yang cukup untuk melakukan aktivitas

Q	Item Penilaian
Q31	Saya cenderung melihat segala hal secara pesimistis
Q32	Saya memiliki kecenderungan merasa sedih dan depresi
Q33	Saya sering kehilangan minat pada berbagai aktivitas sehari-hari
Q34	Saya merasa kurang berharga
Q35	Saya cenderung merasakan hidup tidak bermakna
Q36	Saya tidak bisa menikmati hal-hal yang biasanya menyenangkan
Q37	Saya merasa putus asa dan tanpa harapan
Q38	Saya tidak memiliki ketertarikan pada banyak hal
Q39	Saya merasa tidak berharga atau tidak berarti
Q40	Saya merasa kehilangan harapan terhadap masa depan
Q41	Saya merasa bahwa hidup tidak bermakna
Q42	Saya merasa kesulitan dalam berinisiatif ketika melakukan sesuatu

Berdasarkan pada 42 aspek penilaian, akan dikelompokkan menjadi tiga kategori yaitu skala depresi, kecemasan dan stres antara lain:

- Kategori stres ditunjukkan oleh Q1 – Q14
- Kategori kecemasan ditunjukkan oleh Q15 – Q28
- Kategori depresi ditunjukkan oleh Q29 – Q 42

Pada ketiga kategori diatas memiliki skala penilaian yang dapat dilihat pada Tabel 3.2.

Tabel 3.2 Skala Penilaian

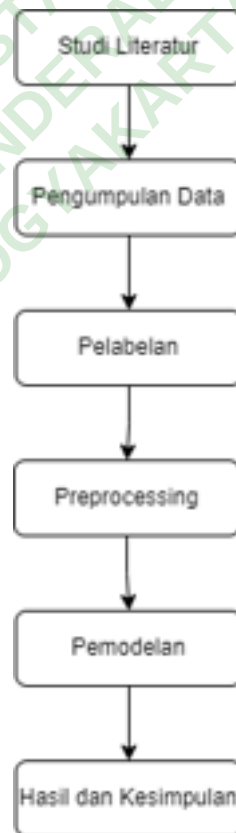
No.	Depresi	Kecemasan	Stres	Total Score (max)
Normal	0 – 9	0 – 7	0 – 14	30
Ringan	10 – 13	8 – 9	15 – 18	40
Sedang	14 – 20	10 – 14	19 – 25	59
Parah	21 – 27	15 – 19	26 – 33	79
Sangat Parah	> 28	> 20	> 34	> 82

Sistem operasi dan program-program aplikasi yang dipergunakan dalam dalam penelitian terkait klasifikasi tingkat kesehatan mental pada mahasiswa akhir studi ini adalah:

1. Sistem Operasi: Windows 10 atau versi baru.
2. Database Manager XAMPP versi 8.2 12-0 atau yang terbaru.
3. Jupyter Notebook
4. Bahasa Pemrograman Python
5. Anaconda
6. Visual Studio Code

3.2 JALAN PENELITIAN

Pada penelitian ini akan dilakukan secara berurutan sebagaimana pada alur penelitian. Jalan penelitian pada penelitian ini dapat dilihat pada Gambar 3.1:



Gambar 3.1 Jalan Penelitian

3.2.1 Pengumpulan Data

Langkah awal dari penelitian ini adalah mengumpulkan *dataset* yang sesuai dengan tujuan penelitian, yaitu *dataset* yang berisi data tingkat kesehatan mental berbasis DASS-42. *Dataset* ini mencakup variabel-variabel yang relevan dengan penelitian menggunakan *Depression, Anxiety, Stress Scales – 42*. *Dataset* yang digunakan pada penelitian ini diperoleh dengan menyebarkan kuesioner Kesehatan mental DASS-42 dengan studi kasus mahasiswa Fakultas Teknik dan Teknologi Informasi di Universitas Jenderal Achmad Yani Yogyakarta. Informasi lebih jelasnya mengenai *dataset* ini dapat dilihat pada Gambar 3.2.

Email address	Name	Student_ID	Gender	Department	Years	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32	G33	G34	G35	G36	G37	G38	G39	G40	G41	G42		
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Gambar 3.2 *Dataset* yang telah diperbaiki

Pada *dataset* diatas terdapat beberapa atribut yang digunakan, beberapa atribut tersebut yakni, diantaranya dapat dilihat pada Tabel 3.3.

Tabel 3.3 Atribut yang digunakan dalam *dataset*

No.	Atribut	Keterangan
1	Email Address	Alamat Email Responden
2	Name	Nama Responden
3	Student_ID	Nomor Induk Mahasiswa / Nomor Pokok Mahasiswa
4	Gender	Jenis Kelamin
5	Department	Program Studi Responden
6	Years	Tahun Angkatan

7	Q1, Q2, Q3,, Q42	Pertanyaan 1, Pertanyaan 2, Pertanyaan 3,, Pertanyaan 42
8	Stress_Factor	Faktor yang melatarbelakangi Kesehatan mahasiswa

3.2.2 Pelabelan Data

Pelabelan data merupakan proses yang dilakukan untuk mengidentifikasi data dengan menambah label pada data tersebut yang kemudian akan dipelajari oleh *machine learning*. Pada penelitian ini, pelabelan dilakukan untuk mengelompokkan data berdasarkan tingkat stres, kecemasan dan depresi. Namun sebelum melakukan pengelompokan tersebut perlu dilakukan pengidentifikasian atribut dari masing-masing kategori, kemudian dikalkulasikan berdasarkan kategori *stress*, kecemasan dan depresi. Dari hasil kalkulasi, setiap nilai pada kategori akan dilakukan pelabelan berdasarkan tingkat normal, ringan, sedang, parah dan sangat parah. Pelabelan ini didasarkan pada skala penilaian pada Tabel 3.2.

Kode pelabelan yang dilakukan di penelitian ini dapat dilihat pada kode berikut.

```
# fungsi klasifikasi dalam menentukan label tingkat depresi,
kecemasan dan stress
def classify_depression(score):
    if score <= 9:
        return 'Normal'
    elif score <= 13:
        return 'Ringan'
    elif score <= 20:
        return 'Sedang'
    elif score <= 27:
        return 'Parah'
    else:
        return 'Sangat Parah'

def classify_anxiety(score):
    if score <= 7:
        return 'Normal'
```

```

elif score <= 9:
    return 'Ringan'
elif score <= 14:
    return 'Sedang'
elif score <= 19:
    return 'Parah'
else:
    return 'Sangat Parah'

def classify_stress(score):
    if score <= 14:
        return 'Normal'
    elif score <= 18:
        return 'Ringan'
    elif score <= 25:
        return 'Sedang'
    elif score <= 33:
        return 'Parah'
    else:
        return 'Sangat Parah'

def classify_mental_health(score):
    if score <= 30:
        return 'Normal'
    elif score <= 40:
        return 'Ringan'
    elif score <= 59:
        return 'Sedang'
    elif score <= 79:
        return 'Parah'
    else:
        return 'Sangat Parah'

```

3.2.3 *Preprocessing Data*

Preprocessing data adalah proses yang dilakukan untuk mempersiapkan dan melakukan transformasi data *raw* menjadi format tertentu yang

dapat memudahkan proses analisis ataupun pemodelan. Pada penelitian ini dilakukan beberapa tahap preprocessing data diantaranya meliputi pengecekan NaN, menghapus kolom, pengecekan data Null, pengecekan tipe data dan mengubah tipe kolom/atribut. Pada tahap pengecekan data NaN dilakukan menggunakan kode berikut ini:

```
# check data NaN
data_baru.isna().sum()
```

Tahap preprocessing selanjutnya adalah menghapus kolom/atribut Stress_Factor. Atribut Stress_Factor merupakan kolom optional yang tidak dibutuhkan dalam penelitian ini. Kode yang digunakan pada tahap ini yaitu menggunakan fungsi drop untuk menghapus kolom, untuk informasi lebih detail dapat dilihat pada kode berikut ini:

```
#drop/menghapus kolom ---> stress factor
data_new = data_baru.drop(['Stress_Factor'], axis=1)
```

Tahap selanjutnya yaitu mengecek data *null*. Data null adalah data yang memiliki nilai kosong atau tidak memiliki nilai. Berikut ini merupakan kode yang digunakan untuk mengecek data *null*:

```
#check data null
data_new.isnull().sum()
```

Kemudian tahap mengubah tipe kolom. Sebelum mengubah tipe kolom perlu dilakukan pengecekan tipe data untuk mengidentifikasi atribut yang membutuhkan pengubahan tipe data yang dapat dilakukan menggunakan kode sebagai berikut.

```
# check type data
data_new.dtypes
```


Selanjutnya yaitu tahap mengubah tipe kolom dilakukan menggunakan “*LabelEncoder*” untuk mengubah tipe data “*object*” menjadi tipe data numerik. Kode yang digunakan untuk mengubah tipe data pada penelitian ini adalah sebagai berikut.

```
#convert kolom bertipe data object menjadi numerik
label_encoder = LabelEncoder()
for column in data_new.columns:
    data_new[column] =
    label_encoder.fit_transform(data_new[column])
```

3.2.4 Pemodelan Data

Pada pemodelan data terdapat beberapa tahap yang harus dilakukan diantaranya sebagai berikut.

1. Menentukan fitur dan target

Pada penelitian ini dilakukan 3 kali pemodelan data dimana pemodelan dilakukan pada kategori depresi, kecemasan dan stres. Fitur adalah variabel independen(X) yang digunakan untuk memprediksi variabel dependen(y). Pada tahap menentukan fitur dan target dapat dilakukan dengan kode berikut ini:

```
# Define features and target
X = data.drop(columns=['Email address', 'Name',
    'Student_ID', 'Depression_Level', 'Anxiety_Level',
    'Stress_Level', 'Total_Depression_Score', 'Total_Anxiety_Score',
    'Total_Stress_Score', 'Total_Score'])
y = data['Stress_Level']
```

melalui kode diatas, *dataset* akan dikelompokkan menjadi fitur(X) dengan menghapus kolom/atribut yang tidak diperlukan dan target(y). kolom yang menjadi target(y) pada kode diatas adalah kolom “*Stress_Level*”. Pemodelan pada kategori kecemasan dan depresi dapat dilakukan menggunakan kode yang sama dengan mengubah target menjadi “*Anxiety_Level*” atau “*Depression_Level*”. Pemodelan ini dilakukan secara terpisah.

2. *Splitting data*

Splitting data adalah tahapan yang dilakukan untuk memisahkan data menjadi data training dan data testing. Berikut ini adalah kode yang digunakan untuk tahapan *splitting data*:

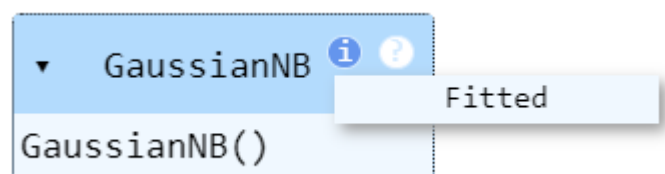
```
# Split the data
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
```

3. Implementasi *Naïve Bayes Classification*

Implementasi *naïve bayes classification* adalah tahapan pembuatan model dengan mengaplikasikan algoritma *naïve bayes classification*. Berikut ini adalah kode yang digunakan untuk pembuatan model *naïve bayes classification*:

```
# Implement Naive Bayes classification
nb = GaussianNB()
nb.fit(X_train_sm, y_train_sm)
```

Tujuan dari proses implementasi *naïve bayes classification* ini adalah untuk membangun sebuah model klasifikasi yang dapat memprediksi kategori dari data baru berdasarkan fitur yang telah dipelajari dari data pelatihan. Output yang dihasilkan ketika menjalankan kode implementasi *naïve bayes classification* yang ditunjukkan pada Gambar 3.3:



Gambar 3.3 Output Implementasi *naïve bayes classification*

Berdasarkan pada Gambar 3.3 dapat dilihat bahwa *dataset* yang digunakan pada penelitian ini bersifat fit dengan model yang digunakan. Dalam mengukur tingkat fit data menggunakan GaussianNB

4. Memprediksi data

Setelah tahap implementasi, langkah selanjutnya adalah membuat prediksi data dengan menerapkan kode seperti berikut.

```
# Predict on test data
y_pred = nb.predict(X_test)
```

3.2.5 Visualisasi Data

Visualisasi data pada penelitian ini dilakukan dengan membuat visualisasi yang berdasarkan pada tingkat stres, kecemasan dan depresi yang dikelompokkan berdasarkan tingkat normal, ringan, sedang, parah, dan sangat parah. Visualisasi data dilakukan secara terpisah berdasarkan kategori stres, kecemasan dan depresi. Kode yang digunakan pada setiap kategori adalah sama, namun dibutuhkan penyesuaian terhadap masing-masing kategori agar menghasilkan visualisasi yang valid. Berikut ini adalah kode yang dapat diterapkan dalam melakukan visualisasi data.

```
# Visualisasi
sns.set(style="whitegrid")
sns.countplot(x='Anxiety_Level', data=data,
order=['Normal', 'Ringan', 'Sedang', 'Parah', 'Sangat Parah'],
hue_order=['Informatika', 'Sistem Informasi', 'Teknik Industri',
'Teknologi Informasi'])
plt.xlabel('Anxiety Level')
plt.ylabel('Count')

# Menambahkan jumlah data di atas setiap batang
for p in plt.gca().patches:
    plt.annotate(f'{int(p.get_height())}', (p.get_x() +
p.get_width() / 2., p.get_height()), ha='center', va='baseline',
fontsize=12, color='black', xytext=(0, 5),textcoords='offset points')

plt.tight_layout()
plt.show()
```